

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US2005/010912

International filing date: 31 March 2005 (31.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US  
Number: 60/559,011  
Filing date: 01 April 2004 (01.04.2004)

Date of receipt at the International Bureau: 12 August 2005 (12.08.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

100218



# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

*August 02, 2005*

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/559,011

FILING DATE: April 01, 2004

RELATED PCT APPLICATION NUMBER: PCT/US05/10912



Certified by

*Don W. Dudas*

Under Secretary of Commerce  
for Intellectual Property  
and Director of the United States  
Patent and Trademark Office

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**PROVISIONAL APPLICATION FOR PATENT COVER SHEET**

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EL 961008135 US

22658 U.S. PTO  
60559011

040104

INVENTOR(S)			
Given Name (first and middle (if any))	Family Name or Surname	Residence (City and either State or Foreign Country)	
Steven Andreas Stefan M.	MAH BRAUN KAMMERER	San Diego, California San Diego, California San Diego, California	
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto			
TITLE OF THE INVENTION (500 characters max)			
METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF			
Direct all correspondence to: CORRESPONDENCE ADDRESS			
<input checked="" type="checkbox"/> Customer Number: <u>25225</u> OR <input type="checkbox"/> Firm or Individual Name			
Address			
City	State	Zip	
Country	Telephone	Fax	
ENCLOSED APPLICATION PARTS (check all that apply)			
<input checked="" type="checkbox"/> Specification Number of Pages	<u>214</u>	<input type="checkbox"/> CD(s), Number	<u>          </u>
<input checked="" type="checkbox"/> Drawing(s) Number of Sheets	<u>2</u>	<input checked="" type="checkbox"/> Other	<u>Return Receipt Postcard</u>
<input checked="" type="checkbox"/> Application Data Sheet. See 37 CFR 1.76 (4 pages) (specify):			
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT			
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.			
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees.			
<input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: <u>03-1952</u>			
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.			
FILING FEE AMOUNT (\$) <u>80.00</u>			
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.			
<input checked="" type="checkbox"/> No <input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are:			

[Page 1 of 2]

Respectfully submitted,

Date April 1, 2004

SIGNATURE

TYPED OR

PRINTED NAME

Bruce D. Grant

TELEPHONE

(858) 720-7962

REGISTRATION NO.

(if appropriate)

47,608

Docket Number:

524593008600**USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT**

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail, Airbill No. EL 961008135 US, in an envelope addressed to: Mail Stop, Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date shown below.

Dated: 4/1/04Signature: Deborah Wykes

(Deborah Wykes)

**PROVISIONAL APPLICATION COVER SHEET**  
*Additional Page*

PTO/SB/16 (08-03)

Approved for use through 07/31/06. OMB 0651-0032

U.S. Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Docket Number		524593008600
INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle (if any))	Family or Surname	Residence (City and either State or Foreign Country)
Matthew Roberts Rikard Henry Maria	NELSON RENELAND LANGDOWN	San Marcos, California San Diego, California San Diego, California

## METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF

### Field of the Invention

[0001] The invention relates to genetic methods for identifying risk of osteoarthritis and treatments that specifically target such diseases.

### Background

[0002] Osteoarthritis (OA) is a chronic disease usually affecting weight-bearing synovial joints. There are approximately 20 million Americans affected by OA and it is the leading cause of disability in the United States. In addition to extensive human suffering, OA also accounts for nearly all knee replacements and more than half of all hip replacements in the United States. Despite its prevalence, OA is poorly understood and there are few treatments available besides anti-inflammatory drugs and joint replacement.

[0003] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke.

[0004] OA is characterized by the breakdown of cartilage in joints. Cartilage in joints cushions the ends of bones, and cartilage breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rheum. 32:241-246 (1989)).

### Summary

[0005] It has been discovered that certain polymorphic variations in human genomic DNA are associated with osteoarthritis. In particular, polymorphic variants in loci containing *KIAA0296*, *chrom 4* and *SNW1* regions and other regions in Table A of human genomic DNA have been associated with risk of osteoarthritis.

[0006] Thus, featured herein are methods for identifying a subject at risk of osteoarthritis and/or a risk of osteoarthritis in a subject, which comprise detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in or around the loci described herein in a human nucleic acid sample. In an embodiment, two or more polymorphic variations are detected in two or more regions of which one is the *KIAA0296*, *chrom 4* or *SNW1* region or other region in Table A. In certain

embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected.

[0007] Also featured are nucleic acids that include one or more polymorphic variations associated with occurrence of osteoarthritis, as well as polypeptides encoded by these nucleic acids. In addition, provided are methods for identifying candidate therapeutic molecules for treating osteoarthritis, as well as methods for treating osteoarthritis in a subject by identifying a subject at risk of osteoarthritis and treating the subject with a suitable prophylactic, treatment or therapeutic molecule.

[0008] Also provided are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A, with a RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid designed from a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A. In an embodiment, the RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid is designed from a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A that includes one or more polymorphic variations associated with osteoarthritis, and in some instances, specifically interacts with such a nucleotide sequence. Further, provided are arrays of nucleic acids bound to a solid surface, in which one or more nucleic acid molecules of the array have a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a fragment or substantially identical nucleic acid thereof, or a complementary nucleic acid of the foregoing. Featured also are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *KIAA0296*, *chrom 4* or *SNW1* polypeptide or other polypeptide referenced in Table A, with an antibody that specifically binds to the polypeptide. In an embodiment, the antibody specifically binds to an epitope in the polypeptide that includes a non-synonymous amino acid modification associated with osteoarthritis (e.g., results in an amino acid substitution in the encoded polypeptide associated with osteoarthritis). In certain embodiments, the antibody selectively binds to an epitope in the *KIAA0296*, *chrom 4* or *SNW1* polypeptide, or other polypeptide referenced in Table A, having an amino acid associated with osteoarthritis. Thus, featured is an antibody that binds an epitope having an amino acid encoded by rs734784, rs1042164, rs749670, rs955592 and/or rs1040461, such as a valine or isoleucine encoded by rs734784 (e.g., a valine at position 489 in a *KCNS1* polypeptide), a valine or alanine encoded by rs1042164 (e.g., a valine at position 133 in a *ETR101* polypeptide), a glutamate or glycine encoded by rs749670 (e.g., a glutamate at position 327 in a *KIAA0296* polypeptide), a threonine or isoleucine encoded by rs955592 (e.g., a threonine at position 70 in a *FLJ21977* polypeptide) or a glycine or serine encoded by rs1040461 (e.g., a glycine at position 207 in a *RAB23* polypeptide) at the corresponding position in the polypeptide.

### Brief Description of the Drawings

[0009] Figures 1A-1C show proximal SNPs in *KIAA0296*, *chrom 4* and *SNW1* regions of genomic DNA, respectively. The position of each SNP in the chromosome is shown on the x-axis and the y-axis provides the negative logarithm of the p-value comparing the estimated allele to that of the control group. Also shown in the figures are exons and introns of the regions in the approximate chromosomal positions.

### Detailed Description

[0010] It has been discovered that a polymorphic variant in a locus containing a *KIAA0296*, *chrom 4* or *SNW1* region is associated with occurrence of osteoarthritis in subjects. Thus, detecting genetic determinants associated with an increased risk of osteoarthritis occurrence can lead to early identification of a predisposition to osteoarthritis and early prescription of preventative measures. Also, associating a *KIAA0296*, *chrom 4* or *SNW1* polymorphic variant and other variants referenced in Table A with osteoarthritis has provided new targets for screening molecules useful in treatments of osteoarthritis.

#### Osteoarthritis and Sample Selection

[0011] Osteoarthritis (OA), or degenerative joint disease, is one of the oldest and most common types of arthritis. It is characterized by the breakdown of the joint's cartilage. Cartilage is the part of the joint that cushions the ends of bones, and its breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rheum. 32:241-246 (1989)).

[0012] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke.

[0013] Osteoarthritis affects an estimated 20.7 million Americans, mostly after age 45, with women more commonly affected than men. Physicians make a diagnosis of OA based on a physical exam and history of symptoms. X-rays are used to confirm diagnosis. Most people over 60 reflect the disease on X-ray, and about one-third have actual symptoms.

[0014] There are many factors that can cause OA. Obesity may lead to osteoarthritis of the knees. In addition, people with joint injuries due to sports, work-related activity or accidents may be at increased risk of developing OA.

[0015] Genetics has a role in the development of OA. Some people may be born with defective cartilage or with slight defects in the way that joints fit together. As a person ages, these defects may

cause early cartilage breakdown in the joint or the inability to repair damaged or deteriorated cartilage in the joint.

[0016] Inclusion or exclusion of samples for an osteoarthritis pool may be based upon the following criteria: ethnicity (e.g., samples derived from an individual characterized as Caucasian); parental ethnicity (e.g., samples derived from an individual of British paternal and maternal descent); relevant phenotype information for the individual (e.g., case samples derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic). Control samples may be selected based on relevant phenotype information for the individual (e.g., derived from individuals free of OA at several sites (knee, hand, hip etc)); and no family history of OA and/or rheumatoid arthritis. Additional phenotype information collected for both cases and controls may include age of the individual, gender, family history of OA, diagnosis with osteoarthritis (joint location of OA, date of primary diagnosis, age of individual as of primary diagnosis), knee history (current symptoms, any major knee injury, meniscectomy, knee replacement surgery, age of surgery), HRT history, osteoporosis diagnosis.

[0017] Based in part upon selection criteria set forth above, individuals having osteoarthritis can be selected for genetic studies. Also, individuals having no history of osteoarthritis often are selected for genetic studies, as described hereafter.

#### Polymorphic Variants Associated with Osteoarthritis

[0018] A genetic analysis provided herein linked osteoarthritis with polymorphic variant nucleic acid sequences in the human genome. As used herein, the term “polymorphic site” refers to a region in a nucleic acid at which two or more alternative nucleotide sequences are observed in a significant number of nucleic acid samples from a population of individuals. A polymorphic site may be a nucleotide sequence of two or more nucleotides, an inserted nucleotide or nucleotide sequence, a deleted nucleotide or nucleotide sequence, or a microsatellite, for example. A polymorphic site that is two or more nucleotides in length may be 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more, 20 or more, 30 or more, 50 or more, 75 or more, 100 or more, 500 or more, or about 1000 nucleotides in length, where all or some of the nucleotide sequences differ within the region. A polymorphic site is often one nucleotide in length, which is referred to herein as a “single nucleotide polymorphism” or a “SNP.”

[0019] Where there are two, three, or four alternative nucleotide sequences at a polymorphic site, each nucleotide sequence is referred to as a “polymorphic variant” or “nucleic acid variant.” Where two polymorphic variants exist, for example, the polymorphic variant represented in a minority of samples from a population is sometimes referred to as a “minor allele” and the polymorphic variant that is more prevalently represented is sometimes referred to as a “major allele.” Many organisms possess a copy of each chromosome (e.g., humans), and those individuals who possess two major alleles or two minor



alleles are often referred to as being “homozygous” with respect to the polymorphism, and those individuals who possess one major allele and one minor allele are normally referred to as being “heterozygous” with respect to the polymorphism. Individuals who are homozygous with respect to one allele are sometimes predisposed to a different phenotype as compared to individuals who are heterozygous or homozygous with respect to another allele.

[0020] In genetic analysis that associate polymorphic variants with osteoarthritis, samples from individuals having osteoarthritis and individuals not having osteoarthritis often are allelotyped and/or genotyped. The term “allelotype” as used herein refers to a process for determining the allele frequency for a polymorphic variant in pooled DNA samples from cases and controls. By pooling DNA from each group, an allele frequency for each SNP in each group is calculated. These allele frequencies are then compared to one another. The term “genotyped” as used herein refers to a process for determining a genotype of one or more individuals, where a “genotype” is a representation of one or more polymorphic variants in a population.

[0021] A genotype or polymorphic variant may be expressed in terms of a “haplotype,” which as used herein refers to two or more polymorphic variants occurring within genomic DNA in a group of individuals within a population. For example, two SNPs may exist within a gene where each SNP position includes a cytosine variation and an adenine variation. Certain individuals in a population may carry one allele (heterozygous) or two alleles (homozygous) having the gene with a cytosine at each SNP position. As the two cytosines corresponding to each SNP in the gene travel together on one or both alleles in these individuals, the individuals can be characterized as having a cytosine/cytosine haplotype with respect to the two SNPs in the gene.

[0022] As used herein, the term “phenotype” refers to a trait which can be compared between individuals, such as presence or absence of a condition, a visually observable difference in appearance between individuals, metabolic variations, physiological variations, variations in the function of biological molecules, and the like. An example of a phenotype is occurrence of osteoarthritis.

[0023] Researchers sometimes report a polymorphic variant in a database without determining whether the variant is represented in a significant fraction of a population. Because a subset of these reported polymorphic variants are not represented in a statistically significant portion of the population, some of them are sequencing errors and/or not biologically relevant. Thus, it is often not known whether a reported polymorphic variant is statistically significant or biologically relevant until the presence of the variant is detected in a population of individuals and the frequency of the variant is determined. Methods for detecting a polymorphic variant in a population are described herein, specifically in Example 2. A polymorphic variant is statistically significant and often biologically relevant if it is represented in 5% or more of a population, sometimes 10% or more, 15% or more, or 20% or more of a population, and often 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, or 50% or more of a population.

[0024] A polymorphic variant may be detected on either or both strands of a double-stranded nucleic acid. Also, a polymorphic variant may be located within an intron or exon of a gene or within a portion of a regulatory region such as a promoter, a 5' untranslated region (UTR), a 3' UTR, and in DNA (e.g., genomic DNA (gDNA) and complementary DNA (cDNA)), RNA (e.g., mRNA, tRNA, and rRNA), or a polypeptide. Polymorphic variations may or may not result in detectable differences in gene expression, polypeptide structure, or polypeptide function.

[0025] It was determined that polymorphic variations associated with an increased risk of osteoarthritis existed in SEQ ID NO: 1-3 or a nucleotide sequence referenced in Table A. In certain embodiments, polymorphic variants at positions rs552, rs12904, rs2282146, rs734784, rs1042164, rs749670, rs955592, rs1143016, rs755248, rs1055055, rs835409, rs927663, rs8162, rs831038, rs33079, rs1710880, rs4472029, rs799570, rs1282730, rs1518875, rs1568694, rs905042, rs1957723, rs794018, rs707723, rs893861, rs1914903, rs2062232, rs26609, rs1370987, rs1012414, rs435903, rs1248, rs703508, rs226465, rs241448, rs763155, rs1040461, rs462832, rs804194, rs1022646, rs1569112, rs805623, rs1019850, rs1599931, AA, rs279941, rs1062230, rs1859911, rs1477261, rs1191119, rs657780, rs1393890, rs1478714, rs868213, rs690115, rs1465501, rs899173, rs10477, rs926393, rs465271, rs13847 and/or rs738658 in the human genome were associated with an increased risk of osteoarthritis, and in specific embodiments, the corresponding allele in the right-most column in Table A for each position is associated with an increased risk of osteoarthritis. In other embodiments polymorphic variants at positions rs734784, rs1042164, rs749670, rs955592 and rs1040461 were associated with an increased risk of osteoarthritis, and in specific embodiments, a valine encoded by rs734784, a valine encoded by rs1042164, a glutamate encoded by rs749670, a threonine encoded by rs955592 and a glycine encoded by rs1040461 were associated with an increased risk of osteoarthritis.

[0026] Polymorphic variants in and around the *KIAA0296* locus were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 1 selected from the group consisting of 247, 1535, 2386, 6440, 9133, 9143, 9471, 13150, 13717, 14466, 15769, 16870, 18545, 18749, 19123, 20736, 21038, 21046, 21050, 21056, 21706, 23170, 25028, 27871, 28070, 31717, 32019, 32318, 33080, 33101, 34236, 34285, 34818, 35168, 37981, 38113, 38117, 38481, 38615, 38944, 39288, 41385, 42136, 42185, 42353, 42434, 44580, 44675, 45739, 46439, 47457, 47735, 50319, 50708, 51185, 53002, 53064, 53637, 55274, 55825, 55986, 56684, 57653, 57659, 57692, 57775, 61313, 61431, 61699, 62906, 63619, 64664, 68452, 69665, 69681, 70091, 74637, 74760, 76523, 78559, 79549, 79882, 81339, 81681, 81696, 83517, 85431, 86332, 87358, 87725, 89052, 90020, 90231, 90284, 90447, 90601, 90724, 92559, 95176, 95195 and 96822. Polymorphic variants at the following positions in SEQ ID NO: 1 in particular were associated with an increased risk of osteoarthritis: 13150, 21046, 23170, 25028, 44580, 62906, 64664 and 83517. In particular, the following polymorphic variants in SEQ ID NO: 1 were associated with risk of osteoarthritis: a guanine at position 13150, a thymine at position 21046, an

adenine at position 23170, an adenine at position 25028, a guanine at position 44580, a guanine at position 62906, a cytosine at position 64664 and a cytosine at position 83517. A polymorphic variant in a *KIAA0296* polypeptide encoded by rs749670 (e.g., a glutamate at position 327 in the polypeptide) also was associated with increased risk of osteoarthritis.

[0027] Polymorphic variants in and around the *chrom 4* locus were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 2 selected from the group consisting of 211, 7217, 7895, 13308, 14279, 17026, 18271, 20417, 21843, 22069, 22145, 22519, 22539, 23236, 23256, 23402, 23499, 23620, 23871, 24136, 25427, 25866, 26541, 26576, 26689, 26720, 27113, 27164, 27186, 28341, 29160, 29844, 30665, 30830, 31061, 31523, 32326, 32346, 32358, 34909, 34975, 35066, 35096, 35375, 36304, 36712, 36770, 37342, 37412, 37884, 38077, 38300, 38301, 41189, 44408, 44493, 44571, 44670, 45219, 45258, 47261, 48473, 48771, 55292, 56479, 56747, 60620, 60688, 61058, 61129, 61577, 61961, 63351, 63926, 65798, 66043, 66044, 66246, 66318, 66547, 71238, 71283, 71492, 72274, 73762, 74209, 75284, 77347, 77589, 78096, 78606, 78862, 79135, 79146, 79456, 79609, 80086, 80119, 80766, 81110, 81269, 81668, 82433, 82559, 83298, 83821, 84121, 84147, 84543, 84554, 84691, 84727, 85678, 86699, 86700, 86792, 86832, 87045, 87140, 87365, 88342, 88498, 88589, 95502, 96968, 97448, 97568 and 98724. Polymorphic variants at the following positions in SEQ ID NO: 2 in particular were associated with an increased risk of osteoarthritis: 23236, 32358, 47261, 48771, 55292, 60688, 72274, 74209, 77589, 79135, 79456, 79609, 80119, 80766, 81110, 82433, 84121, 84147, 85678, 86699, 86832, 87140 and 88589, where specific embodiments are directed to a polymorphic variant at position 32358, 47261, 74209 and/or 79456. In particular, the following polymorphic variants in SEQ ID NO: 2 were associated with risk of osteoarthritis: an adenine at position 23236, a cytosine at position 32358, a guanine at position 47261, a guanine at position 48771, a cytosine at position 55292, an adenine at position 60688, a guanine at position 72274, a guanine at position 74209, a cytosine at position 77589, an adenine at position 79135, a thymine at position 79456, an adenine at position 79609, an adenine at position 80119, a cytosine at position 80766, an adenine at position 81110, a cytosine at position 82433, a cytosine at position 84121, a thymine at position 84147, a cytosine at position 85678, a thymine at position 86699, an adenine at position 86832, a guanine at position 87140 and an adenine at position 88589.

[0028] Polymorphic variants in and around the *SNNI* locus were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 3 selected from the group consisting of 218, 1440, 1442, 2611, 4317, 4724, 4788, 5202, 5780, 5974, 6644, 7430, 7938, 8095, 8183, 8312, 8352, 9348, 9378, 9617, 9727, 9834, 9899, 10211, 10377, 10695, 10729, 10730, 11433, 11951, 12697, 12982, 14419, 14501, 14983, 15280, 15475, 15888, 15976, 16307, 16442, 17255, 18948, 19435, 19753, 20021, 20022, 20503, 20590, 21804, 21919, 21990, 22412, 22536, 23432, 23468, 23772, 24325, 24773, 26274, 27440, 28561, 30071, 31764, 33008, 35310, 35460, 37112, 37285, 37747, 38057, 38859,

38860, 39525, 40216, 40281, 41453, 42091, 42513, 42935, 42985, 43003, 43281, 43716, 43866, 44234, 44596, 44871, 45005, 45282, 47178, 47816, 47887, 48134, 48135, 48276, 48400, 48798, 48803, 49146, 49969, 51059, 51064, 53285, 54560, 54748, 54785, 55102, 55644, 55705, 55841, 56623, 56825, 56827, 56892, 59150, 59958, 60231, 60524, 61871, 62226, 63230, 63468, 63787, 65732, 65989, 68832, 69904, 70365, 70886, 73088, 73103, 75934, 75966, 76273, 77943, 78466, 78861, 78872, 79836, 80908, 81509, 83576, 83662, 83782, 84282, 84444, 85129, 85151, 85296, 85809, 86387, 86494, 89786, 89894, 90122, 92067, 92187, 92312, 92824, 93733, 96553 and 96941. Polymorphic variants at the following positions in SEQ ID NO: 3 in particular were associated with an increased risk of osteoarthritis: 4788, 8312, 9378, 9727, 9899, 10211, 27440, 40216, 40281, 42091, 43866, 48803, 51059, 55644, 56623, 73103, 78872, 79836, 85129, 92824 and 96941. In particular, the following polymorphic variants in SEQ ID NO: 3 were associated with risk of osteoarthritis: a guanine at position 4788, a thymine at position 8312, a deletion at position 9378, a cytosine at position 9727, a guanine at position 9899, a cytosine at position 10211, a guanine at position 27440, a guanine at position 40216, a cytosine at position 40281, an adenine at position 42091, a guanine at position 43866, an adenine at position 48803, an adenine at position 51059, an adenine at position 55644, a cytosine at position 56623, a cytosine at position 73103, an adenine at position 78872, a guanine at position 79836, a cytosine at position 85129, a guanine at position 92824 and an adenine at position 96941.

[0029] Based in part upon analyses summarized in Figures 1A-1C, regions with significant association have been identified in regions associated with osteoarthritis. Any polymorphic variants associated with osteoarthritis in a region of significant association can be utilized for embodiments described herein. For example, polymorphic variants in a region spanning chromosome positions 31118000 to 31129000 (approximately 11,000 nucleotides in length) in a *KLAA0296* locus, a region spanning chromosome positions 36914000 to 36931000 (approximately 17,000 nucleotides in length) in a *chrom 4* locus and a region spanning chromosome positions 76196500 to 76221500 (approximately 25,000 nucleotides in length) in a *SNW1* locus have significant association (chromosome positions are within NCBI's Genome build 34).

#### Additional Polymorphic Variants Associated with Osteoarthritis

[0030] Also provided is a method for identifying polymorphic variants proximal to an incident, founder polymorphic variant associated with osteoarthritis. Thus, featured herein are methods for identifying a polymorphic variation associated with osteoarthritis that is proximal to an incident polymorphic variation associated with osteoarthritis, which comprises identifying a polymorphic variant proximal to the incident polymorphic variant associated with osteoarthritis, where the incident polymorphic variant is in a *KLAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A. The nucleotide sequence often comprises a polynucleotide sequence

selected from the group consisting of (a) a polynucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (b) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence encoded by a polynucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; and (c) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A or a polynucleotide sequence 90% or more identical to the polynucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A. The presence or absence of an association of the proximal polymorphic variant with osteoarthritis then is determined using a known association method, such as a method described in the Examples hereafter. In an embodiment, the incident polymorphic variant is a polymorphic variant associated with osteoarthritis described herein. In another embodiment, the proximal polymorphic variant identified sometimes is a publicly disclosed polymorphic variant, which for example, sometimes is published in a publicly available database. In other embodiments, the polymorphic variant identified is not publicly disclosed and is discovered using a known method, including, but not limited to, sequencing a region surrounding the incident polymorphic variant in a group of nucleic samples. Thus, multiple polymorphic variants proximal to an incident polymorphic variant are associated with osteoarthritis using this method.

[0031] The proximal polymorphic variant often is identified in a region surrounding the incident polymorphic variant. In certain embodiments, this surrounding region is about 50 kb flanking the first polymorphic variant (*e.g.* about 50 kb 5' of the first polymorphic variant and about 50 kb 3' of the first polymorphic variant), and the region sometimes is composed of shorter flanking sequences, such as flanking sequences of about 40 kb, about 30 kb, about 25 kb, about 20 kb, about 15 kb, about 10 kb, about 7 kb, about 5 kb, or about 2 kb 5' and 3' of the incident polymorphic variant. In other embodiments, the region is composed of longer flanking sequences, such as flanking sequences of about 55 kb, about 60 kb, about 65 kb, about 70 kb, about 75 kb, about 80 kb, about 85 kb, about 90 kb, about 95 kb, or about 100 kb 5' and 3' of the incident polymorphic variant.

[0032] In certain embodiments, polymorphic variants associated with osteoarthritis are identified iteratively. For example, a first proximal polymorphic variant is associated with osteoarthritis using the methods described above and then another polymorphic variant proximal to the first proximal polymorphic variant is identified (*e.g.*, publicly disclosed or discovered) and the presence or absence of an association of one or more other polymorphic variants proximal to the first proximal polymorphic variant with osteoarthritis is determined.

[0033] The methods described herein are useful for identifying or discovering additional polymorphic variants that may be used to further characterize a gene, region or loci associated with a condition, a disease (*e.g.*, osteoarthritis), or a disorder. For example, allelotyping or genotyping data from the additional polymorphic variants may be used to identify a functional mutation or a region of

linkage disequilibrium. In certain embodiments, polymorphic variants identified or discovered within a region comprising the first polymorphic variant associated with osteoarthritis are genotyped using the genetic methods and sample selection techniques described herein, and it can be determined whether those polymorphic variants are in linkage disequilibrium with the first polymorphic variant. The size of the region in linkage disequilibrium with the first polymorphic variant also can be assessed using these genotyping methods. Thus, provided herein are methods for determining whether a polymorphic variant is in linkage disequilibrium with a first polymorphic variant associated with osteoarthritis, and such information can be used in prognosis/diagnosis methods described herein.

#### Isolated Nucleic Acids

[0034] Featured herein are isolated *KIAA0296*, *chrom 4* or *SNW1* nucleic acid variants depicted in SEQ ID NO: 1-3 or referenced in Table A, and substantially identical nucleic acids thereof. A nucleic acid variant may be represented on one or both strands in a double-stranded nucleic acid or on one chromosomal complement (heterozygous) or both chromosomal complements (homozygous).

[0035] As used herein, the term “nucleic acid” includes DNA molecules (*e.g.*, a complementary DNA (cDNA) and genomic DNA (gDNA)) and RNA molecules (*e.g.*, mRNA, rRNA, siRNA and tRNA) and analogs of DNA or RNA, for example, by use of nucleotide analogs. The nucleic acid molecule can be single-stranded and it is often double-stranded. The term “isolated or purified nucleic acid” refers to nucleic acids that are separated from other nucleic acids present in the natural source of the nucleic acid. For example, with regard to genomic DNA, the term “isolated” includes nucleic acids which are separated from the chromosome with which the genomic DNA is naturally associated. An “isolated” nucleic acid is often free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and/or 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of 5' and/or 3' nucleotide sequences which flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an “isolated” nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. As used herein, the term “gene” refers to a nucleotide sequence that encodes a polypeptide.

[0036] Also included herein are nucleic acid fragments. These fragments often have a nucleotide sequence identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, a nucleotide sequence substantially identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, or a nucleotide sequence that is complementary to the foregoing. The nucleic acid fragment may be identical, substantially identical or homologous to a nucleotide sequence in an exon or an intron in a

nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, and may encode a domain or part of a domain of a polypeptide. Sometimes, the fragment will comprises one or more of the polymorphic variations described herein as being associated with osteoarthritis. The nucleic acid fragment is often 50, 100, or 200 or fewer base pairs in length, and is sometimes about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 2000, 3000, 4000, 5000, 10000, 15000, or 20000 base pairs in length. A nucleic acid fragment that is complementary to a nucleotide sequence identical or substantially identical to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A and hybridizes to such a nucleotide sequence under stringent conditions is often referred to as a "probe." Nucleic acid fragments often include one or more polymorphic sites, or sometimes have an end that is adjacent to a polymorphic site as described hereafter.

[0037] An example of a nucleic acid fragment is an oligonucleotide. As used herein, the term "oligonucleotide" refers to a nucleic acid comprising about 8 to about 50 covalently linked nucleotides, often comprising from about 8 to about 35 nucleotides, and more often from about 10 to about 25 nucleotides. The backbone and nucleotides within an oligonucleotide may be the same as those of naturally occurring nucleic acids, or analogs or derivatives of naturally occurring nucleic acids, provided that oligonucleotides having such analogs or derivatives retain the ability to hybridize specifically to a nucleic acid comprising a targeted polymorphism. Oligonucleotides described herein may be used as hybridization probes or as components of prognostic or diagnostic assays, for example, as described herein.

[0038] Oligonucleotides are typically synthesized using standard methods and equipment, such as the ABI™3900 High Throughput DNA Synthesizer and the EXPEDITE™ 8909 Nucleic Acid Synthesizer, both of which are available from Applied Biosystems (Foster City, CA). Analogs and derivatives are exemplified in U.S. Pat. Nos. 4,469,863; 5,536,821; 5,541,306; 5,637,683; 5,637,684; 5,700,922; 5,717,083; 5,719,262; 5,739,308; 5,773,601; 5,886,165; 5,929,226; 5,977,296; 6,140,482; WO 00/56746; WO 01/14398, and related publications. Methods for synthesizing oligonucleotides comprising such analogs or derivatives are disclosed, for example, in the patent publications cited above and in U.S. Pat. Nos. 5,614,622; 5,739,314; 5,955,599; 5,962,674; 6,117,992; in WO 00/75372; and in related publications.

[0039] Oligonucleotides may also be linked to a second moiety. The second moiety may be an additional nucleotide sequence such as a tail sequence (e.g., a polyadenosine tail), an adapter sequence (e.g., phage M13 universal tail sequence), and others. Alternatively, the second moiety may be a non-nucleotide moiety such as a moiety which facilitates linkage to a solid support or a label to facilitate detection of the oligonucleotide. Such labels include, without limitation, a radioactive label, a fluorescent label, a chemiluminescent label, a paramagnetic label, and the like. The second moiety may

be attached to any position of the oligonucleotide, provided the oligonucleotide can hybridize to the nucleic acid comprising the polymorphism.

#### Uses for Nucleic Acid Sequence

[0040] Nucleic acid coding sequences may be used for diagnostic purposes for detection and control of polypeptide expression. Also, included herein are oligonucleotide sequences such as antisense RNA, small-interfering RNA (siRNA) and DNA molecules and ribozymes that function to inhibit translation of a polypeptide. Antisense techniques and RNA interference techniques are known in the art and are described herein.

[0041] Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, hammerhead motif ribozyme molecules may be engineered that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences corresponding to or complementary to *KLAA0296, chrom 4* or *SNW1* nucleotide sequences or other nucleotide sequences referenced in Table A. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between fifteen (15) and twenty (20) ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

[0042] Antisense RNA and DNA molecules, siRNA and ribozymes may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides well known in the art such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

[0043] DNA encoding a polypeptide also may have a number of uses for the diagnosis of diseases, including osteoarthritis, resulting from aberrant expression of a target gene described herein. For example, the nucleic acid sequence may be used in hybridization assays of biopsies or autopsies to diagnose abnormalities of expression or function (*e.g.*, Southern or Northern blot analysis, *in situ* hybridization assays).



[0044] In addition, the expression of a polypeptide during embryonic development may also be determined using nucleic acid encoding the polypeptide. As addressed, *infra*, production of functionally impaired polypeptide is the cause of various disease states, such as osteoarthritis. *In situ* hybridizations using polypeptide as a probe may be employed to predict problems related to osteoarthritis. Further, as indicated, *infra*, administration of human active polypeptide, recombinantly produced as described herein, may be used to treat disease states related to functionally impaired polypeptide. Alternatively, gene therapy approaches may be employed to remedy deficiencies of functional polypeptide or to replace or compete with dysfunctional polypeptide.

#### Expression Vectors, Host Cells, and Genetically Engineered Cells

[0045] Provided herein are nucleic acid vectors, often expression vectors, which contain a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical sequence thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid, or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors may include replication defective retroviruses, adenoviruses and adeno-associated viruses for example.

[0046] A vector can include a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A in a form suitable for expression of an encoded target polypeptide or target nucleic acid in a host cell. A "target polypeptide" is a polypeptide encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical nucleotide sequence thereof. The recombinant expression vector typically includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term "regulatory sequence" includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or inducible sequences. The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, and the like. Expression vectors can be introduced into host cells to produce target polypeptides, including fusion polypeptides.

[0047] Recombinant expression vectors can be designed for expression of target polypeptides in prokaryotic or eukaryotic cells. For example, target polypeptides can be expressed in *E. coli*, insect cells (*e.g.*, using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

[0048] Expression of polypeptides in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino terminus of the recombinant polypeptide. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant polypeptide; 2) to increase the solubility of the recombinant polypeptide; and 3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide to enable separation of the recombinant polypeptide from the fusion moiety subsequent to purification of the fusion polypeptide. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith & Johnson, *Gene* 67: 31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding polypeptide, or polypeptide A, respectively, to the target recombinant polypeptide.

[0049] Purified fusion polypeptides can be used in screening assays and to generate antibodies specific for target polypeptides. In a therapeutic embodiment, fusion polypeptide expressed in a retroviral expression vector is used to infect bone marrow cells that are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six (6) weeks).

[0050] Expressing the polypeptide in host bacteria with an impaired capacity to proteolytically cleave the recombinant polypeptide is often used to maximize recombinant polypeptide expression (Gottesman, S., *Gene Expression Technology: Methods in Enzymology*, Academic Press, San Diego, California 185: 119-128 (1990)). Another strategy is to alter the nucleotide sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, *Nucleic Acids Res.* 20: 2111-2118 (1992)). Such alteration of nucleotide sequences can be carried out by standard DNA synthesis techniques.

[0051] When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. Recombinant mammalian expression vectors are often capable of directing expression of the nucleic acid in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Non-limiting examples of suitable tissue-specific promoters include an albumin promoter (liver-specific; Pinkert *et al.*, *Genes Dev.* 1: 268-277 (1987)), lymphoid-specific promoters (Calame & Eaton, *Adv. Immunol.* 43: 235-275 (1988)), promoters of T cell receptors (Winoto & Baltimore, *EMBO J.* 8: 729-733 (1989)) promoters of immunoglobulins (Banerji *et al.*, *Cell* 33: 729-740 (1983); Queen & Baltimore, *Cell* 33: 741-748 (1983)), neuron-specific

promoters (e.g., the neurofilament promoter; Byrne & Ruddle, *Proc. Natl. Acad. Sci. USA* 86: 5473-5477 (1989)), pancreas-specific promoters (Edlund *et al.*, *Science* 230: 912-916 (1985)), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are sometimes utilized, for example, the murine hox promoters (Kessel & Gruss, *Science* 249: 374-379 (1990)) and the  $\alpha$ -fetopolypeptide promoter (Campe & Tilghman, *Genes Dev.* 3: 537-546 (1989)).

[0052] A *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A also may be cloned into an expression vector in an antisense orientation. Regulatory sequences (e.g., viral promoters and/or enhancers) operatively linked to a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A cloned in the antisense orientation can be chosen for directing constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of cell types. Antisense expression vectors can be in the form of a recombinant plasmid, phagemid or attenuated virus. For a discussion of the regulation of gene expression using antisense genes see, e.g., Weintraub *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) (1986).

[0053] Also provided herein are host cells that include a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A within a recombinant expression vector or a fragment of such a nucleotide sequence which facilitate homologous recombination into a specific site of the host cell genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. Such terms refer not only to the particular subject cell but rather also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, a target polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

[0054] Vectors can be introduced into host cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, transduction/infection, DEAE-dextran-mediated transfection, lipofection, or electroporation.

[0055] A host cell provided herein can be used to produce (*i.e.*, express) a target polypeptide or a substantially identical polypeptide thereof. Accordingly, further provided are methods for producing a target polypeptide using host cells described herein. In one embodiment, the method includes culturing host cells into which a recombinant expression vector encoding a target polypeptide has been introduced

in a suitable medium such that a target polypeptide is produced. In another embodiment, the method further includes isolating a target polypeptide from the medium or the host cell.

[0056] Also provided are cells or purified preparations of cells which include a *KIAA0296, chrom 4* or *SNW1* transgene, or other transgene in Table A, or which otherwise misexpress target polypeptide. Cell preparations can consist of human or non-human cells, e.g., rodent cells, e.g., mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a *KIAA0296, chrom 4* or *SNW1* transgene or other transgene referenced in Table A (e.g., a heterologous form of a *KIAA0296, chrom 4* or *SNW1* gene or other gene referenced in Table A, such as a human gene expressed in non-human cells). The transgene can be misexpressed, e.g., overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous target polypeptide (e.g., expression of a gene is disrupted, also known as a knockout). Such cells can serve as a model for studying disorders which are related to mutated or mis-expressed alleles or for use in drug screening. Also provided are human cells (e.g., a hematopoietic stem cells) transfected with a *KIAA0296, chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A.

[0057] Also provided are cells or a purified preparation thereof (e.g., human cells) in which an endogenous *KIAA0296, chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A is under the control of a regulatory sequence that does not normally control the expression of the endogenous gene. The expression characteristics of an endogenous gene within a cell (e.g., a cell line or microorganism) can be modified by inserting a heterologous DNA regulatory element into the genome of the cell such that the inserted regulatory element is operably linked to the corresponding endogenous gene. For example, an endogenous corresponding gene (e.g., a gene which is "transcriptionally silent," not normally expressed, or expressed only at very low levels) may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell. Techniques such as targeted homologous recombinations, can be used to insert the heterologous DNA as described in, e.g., Chappel, US 5,272,071; WO 91/06667, published on May 16, 1991.

#### Transgenic Animals

[0058] Non-human transgenic animals that express a heterologous target polypeptide (e.g., expressed from a *KIAA0296, chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A, or substantially identical sequence thereof) can be generated. Such animals are useful for studying the function and/or activity of a target polypeptide and for identifying and/or evaluating modulators of the activity of *KIAA0296, chrom 4* or *SNW1* nucleic acids, other nucleic acids referenced in Table A, and encoded polypeptides. As used herein, a "transgenic animal" is a non-human animal such as a mammal (e.g., a non-human primate such as chimpanzee, baboon, or macaque; an ungulate such as an equine, bovine, or caprine; or a rodent such as a rat, a mouse, or an Israeli sand rat), a bird (e.g., a chicken or a

turkey), an amphibian (e.g., a frog, salamander, or newt), or an insect (e.g., *Drosophila melanogaster*), in which one or more of the cells of the animal includes a transgene. A transgene is exogenous DNA or a rearrangement (e.g., a deletion of endogenous chromosomal DNA) that is often integrated into or occurs in the genome of cells in a transgenic animal. A transgene can direct expression of an encoded gene product in one or more cell types or tissues of the transgenic animal, and other transgenes can reduce expression (e.g., a knockout). Thus, a transgenic animal can be one in which an endogenous nucleic acid homologous to a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal (e.g., an embryonic cell of the animal) prior to development of the animal.

[0059] Intronic sequences and polyadenylation signals can also be included in the transgene to increase expression efficiency of the transgene. One or more tissue-specific regulatory sequences can be operably linked to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A to direct expression of an encoded polypeptide to particular cells. A transgenic founder animal can be identified based upon the presence of a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A in its genome and/or expression of encoded mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A can further be bred to other transgenic animals carrying other transgenes.

[0060] Target polypeptides can be expressed in transgenic animals or plants by introducing, for example, a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A into the genome of an animal that encodes the target polypeptide. In preferred embodiments the nucleic acid is placed under the control of a tissue specific promoter, e.g., a milk or egg specific promoter, and recovered from the milk or eggs produced by the animal. Also included is a population of cells from a transgenic animal.

#### Target Polypeptides

[0061] Also featured herein are isolated target polypeptides, which are encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or a nucleotide sequence referenced in Table A (e.g., SEQ ID NO: 1-3 or a sequence referenced in Table A), or a substantially identical nucleotide sequence thereof. The term "polypeptide" as used herein includes proteins and peptides. An "isolated" or "purified" polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. In one embodiment, the language "substantially free"

means preparation of a target polypeptide having less than about 30%, 20%, 10% and more preferably 5% (by dry weight), of non-target polypeptide (also referred to herein as a “contaminating protein”), or of chemical precursors or non-target chemicals. When the target polypeptide or a biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, specifically, where culture medium represents less than about 20%, sometimes less than about 10%, and often less than about 5% of the volume of the polypeptide preparation. Isolated or purified target polypeptide preparations are sometimes 0.01 milligrams or more or 0.1 milligrams or more, and often 1.0 milligrams or more and 10 milligrams or more in dry weight.

[0062] Further included herein are target polypeptide fragments. The polypeptide fragment may be a domain or part of a domain of a target polypeptide. The polypeptide fragment may have increased, decreased or unexpected biological activity. The polypeptide fragment is often 50 or fewer, 100 or fewer, or 200 or fewer amino acids in length, and is sometimes 300, 400, 500, 600, 700, or 900 or fewer amino acids in length. Specific embodiments are directed to a *PTPNI* polypeptide fragment (e.g., rs2282146 in Table A), such as a catalytic domain starting at about amino acid 3 and ending at about amino acid 279. Other embodiments are directed to a *KCNS1* polypeptide fragment (e.g., rs734784 in Table A), such as a voltage gated potassium ion channel domain (e.g., starting at about amino acid 21 and ending at about amino acid 509), a potassium channel tetramerization domain (e.g., starting at about amino acid 52 and ending at about amino acid 155) or an ion transport protein domain (e.g., starting at about amino acid 271 and ending at about amino acid 456), for example. Certain embodiments are directed to a *ANXA6* polypeptide fragment (e.g., rs1012414 in Table A), such as an annexin domain starting at about amino acid 5 and ending at about amino acid 325, an annexin domain starting at about amino acid 179 and ending at about amino acid 507, or an annexin domain starting at about amino acid 355 and ending at about amino acid 673 in isoform 1 or isoform 2 (e.g., an isoform 1 sequence can be accessed using accession number NP\_001146 and an isoform 2 sequence can be accessed using accession number NP\_004024; isoform 2 lacks exon 21 and encodes a protein isoform lacking the six amino acids VAAEIL). Amino acid sequences can be accessed using information in Table A.

[0063] Substantially identical target polypeptides may depart from the amino acid sequences of target polypeptides in different manners. For example, conservative amino acid modifications may be introduced at one or more positions in the amino acid sequences of target polypeptides. A “conservative amino acid substitution” is one in which the amino acid is replaced by another amino acid having a similar structure and/or chemical function. Families of amino acid residues having similar structures and functions are well known. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side

chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Also, essential and non-essential amino acids may be replaced. A “non-essential” amino acid is one that can be altered without abolishing or substantially altering the biological function of a target polypeptide, whereas altering an “essential” amino acid abolishes or substantially alters the biological function of a target polypeptide. Amino acids that are conserved among target polypeptides are typically essential amino acids. In certain embodiments, the polypeptide includes one or more non-synonymous polymorphic variants associated with osteoarthritis, as described above (e.g., a valine encoded by rs734784, a valine encoded by rs1042164, a glutamate encoded by rs749670, a threonine encoded by rs955592 and a glycine encoded by rs1040461).

[0064] Also, target polypeptides may exist as chimeric or fusion polypeptides. As used herein, a target “chimeric polypeptide” or target “fusion polypeptide” includes a target polypeptide linked to a non-target polypeptide. A “non-target polypeptide” refers to a polypeptide having an amino acid sequence corresponding to a polypeptide which is not substantially identical to the target polypeptide, which includes, for example, a polypeptide that is different from the target polypeptide and derived from the same or a different organism. The target polypeptide in the fusion polypeptide can correspond to an entire or nearly entire target polypeptide or a fragment thereof. The non-target polypeptide can be fused to the N-terminus or C-terminus of the target polypeptide.

[0065] Fusion polypeptides can include a moiety having high affinity for a ligand. For example, the fusion polypeptide can be a GST-target fusion polypeptide in which the target sequences are fused to the C-terminus of the GST sequences, or a polyhistidine-target fusion polypeptide in which the target polypeptide is fused at the N- or C-terminus to a string of histidine residues. Such fusion polypeptides can facilitate purification of recombinant target polypeptide. Expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide), and a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A, or a substantially identical nucleotide sequence thereof, can be cloned into an expression vector such that the fusion moiety is linked in-frame to the target polypeptide. Further, the fusion polypeptide can be a target polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression, secretion, cellular internalization, and cellular localization of a target polypeptide can be increased through use of a heterologous signal sequence. Fusion polypeptides can also include all or a part of a serum polypeptide (e.g., an IgG constant region or human serum albumin).

[0066] Target polypeptides can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. Administration of these target polypeptides can be used to affect the bioavailability of a substrate of the target polypeptide and may effectively increase target polypeptide biological activity in a cell. Target fusion polypeptides may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a target polypeptide; (ii) mis-

regulation of the gene encoding the target polypeptide; and (iii) aberrant post-translational modification of a target polypeptide. Also, target polypeptides can be used as immunogens to produce anti-target antibodies in a subject, to purify target polypeptide ligands or binding partners, and in screening assays to identify molecules which inhibit or enhance the interaction of a target polypeptide with a substrate.

[0067] In addition, polypeptides can be chemically synthesized using techniques known in the art (See, *e.g.*, Creighton, 1983 Proteins. New York, N.Y.: W. H. Freeman and Company; and Hunkapiller et al., (1984) Nature July 12 -18;310(5973):105-11). For example, a relative short fragment can be synthesized by use of a peptide synthesizer. Furthermore, if desired, non-classical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the fragment sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid,  $\alpha$ -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid,  $\gamma$ -Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoroamino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

[0068] Polypeptides and polypeptide fragments sometimes are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH<sub>4</sub>; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; and the like. Additional post-translational modifications include, for example, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The polypeptide fragments may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the polypeptide.

[0069] Also provided are chemically modified derivatives of polypeptides that can provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (*see e.g.*, U.S. Pat. No: 4,179,337. The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.



[0070] The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

[0071] The polymers should be attached to the polypeptide with consideration of effects on functional or antigenic domains of the polypeptide. There are a number of attachment methods available to those skilled in the art (e.g., EP 0 401 384 (coupling PEG to G-CSF) and Malik et al. (1992) Exp Hematol. September;20(8):1028-35 (pegylation of GM-CSF using tresyl chloride)). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues, glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. For therapeutic purposes, the attachment sometimes is at an amino group, such as attachment at the N-terminus or lysine group.

[0072] Proteins can be chemically modified at the N-terminus. Using polyethylene glycol as an illustration of such a composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, and the like), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus may be accomplished by reductive alkylation, which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

#### Substantially Identical Nucleic Acids and Polypeptides

[0073] Nucleotide sequences and polypeptide sequences that are substantially identical to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A and the target polypeptide sequences encoded by those nucleotide sequences, respectively, are included

herein. The term "substantially identical" as used herein refers to two or more nucleic acids or polypeptides sharing one or more identical nucleotide sequences or polypeptide sequences, respectively. Included are nucleotide sequences or polypeptide sequences that are 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more (each often within a 1%, 2%, 3% or 4% variability) identical to a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence, or other nucleotide sequence referenced in Table A, or the encoded target polypeptide amino acid sequences. One test for determining whether two nucleic acids are substantially identical is to determine the percent of identical nucleotide sequences or polypeptide sequences shared between the nucleic acids or polypeptides.

[0074] Calculations of sequence identity are often performed as follows. Sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The length of a reference sequence aligned for comparison purposes is sometimes 30% or more, 40% or more, 50% or more, often 60% or more, and more often 70% or more, 80% or more, 90% or more, or 100% of the length of the reference sequence. The nucleotides or amino acids at corresponding nucleotide or polypeptide positions, respectively, are then compared among the two sequences. When a position in the first sequence is occupied by the same nucleotide or amino acid as the corresponding position in the second sequence, the nucleotides or amino acids are deemed to be identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, introduced for optimal alignment of the two sequences.

[0075] Comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. Percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of Meyers & Miller, *CABIOS* 4: 11-17 (1989), which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. Also, percent identity between two amino acid sequences can be determined using the Needleman & Wunsch, *J. Mol. Biol.* 48: 444-453 (1970) algorithm which has been incorporated into the GAP program in the GCG software package (available at the http address [www.gcg.com](http://www.gcg.com)), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. Percent identity between two nucleotide sequences can be determined using the GAP program in the GCG software package (available at http address [www.gcg.com](http://www.gcg.com)), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A set of parameters often used is a Blossum 62 scoring matrix with a gap open penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

[0076] Another manner for determining if two nucleic acids are substantially identical is to assess whether a polynucleotide homologous to one nucleic acid will hybridize to the other nucleic acid under stringent conditions. As use herein, the term "stringent conditions" refers to conditions for hybridization and washing. Stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y., 6.3.1-6.3.6 (1989). Aqueous and non-aqueous methods are described in that reference and either can be used. An example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50°C. Another example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 55°C. A further example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C. Often, stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C. More often, stringency conditions are 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C.

[0077] An example of a substantially identical nucleotide sequence to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A is one that has a different nucleotide sequence but still encodes the same polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A. Another example is a nucleotide sequence that encodes a polypeptide having a polypeptide sequence that is more than 70% or more identical to, sometimes more than 75% or more, 80% or more, or 85% or more identical to, and often more than 90% or more and 95% or more identical to a polypeptide sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A.

[0078] Nucleotide sequences in SEQ ID NO: 1-3 or referenced in Table A and amino acid sequences of encoded polypeptides can be used as "query sequences" to perform a search against public databases to identify other family members or related sequences, for example. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al.*, *J. Mol. Biol.* 215: 403-10 (1990). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleotide sequences in SEQ ID NO: 1-3 or referenced in Table A. BLAST polypeptide searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to polypeptides encoded by the nucleotide sequences of SEQ ID NO: 1-3 or referenced in Table A. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, *Nucleic Acids Res.* 25(17): 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, default parameters of

the respective programs (*e.g.*, XBLAST and NBLAST) can be used (*see* the http address [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)).

[0079] A nucleic acid that is substantially identical to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A may include polymorphic sites at positions equivalent to those described herein when the sequences are aligned. For example, using the alignment procedures described herein, SNPs in a sequence substantially identical to a sequence in SEQ ID NO: 1-3 or referenced in Table A can be identified at nucleotide positions that match (*i.e.*, align) with nucleotides at SNP positions in each nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A. Also, where a polymorphic variation results in an insertion or deletion, insertion or deletion of a nucleotide sequence from a reference sequence can change the relative positions of other polymorphic sites in the nucleotide sequence.

[0080] Substantially identical nucleotide and polypeptide sequences include those that are naturally occurring, such as allelic variants (same locus), splice variants, homologs (different locus), and orthologs (different organism) or can be non-naturally occurring. Non-naturally occurring variants can be generated by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product). Orthologs, homologs, allelic variants, and splice variants can be identified using methods known in the art. These variants normally comprise a nucleotide sequence encoding a polypeptide that is 50% or more, about 55% or more, often about 70-75% or more or about 80-85% or more, and sometimes about 90-95% or more identical to the amino acid sequences of target polypeptides or a fragment thereof. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A or a fragment of this sequence. Nucleic acid molecules corresponding to orthologs, homologs, and allelic variants of a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A can further be identified by mapping the sequence to the same chromosome or locus as the nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A.

[0081] Also, substantially identical nucleotide sequences may include codons that are altered with respect to the naturally occurring sequence for enhancing expression of a target polypeptide in a particular expression system. For example, the nucleic acid can be one in which one or more codons are altered, and often 10% or more or 20% or more of the codons are altered for optimized expression in bacteria (*e.g.*, *E. coli*), yeast (*e.g.*, *S. cerevisiae*), human (*e.g.*, 293 cells), insect, or rodent (*e.g.*, hamster) cells.

Methods for Identifying Risk of Osteoarthritis

[0082] Methods for prognosing and diagnosing osteoarthritis are included herein. These methods include detecting the presence or absence of one or more polymorphic variations in a nucleotide sequence associated with osteoarthritis, such as variants in or around the loci set forth herein, or a substantially identical sequence thereof, in a sample from a subject, where the presence of a polymorphic variant described herein is indicative of a risk of osteoarthritis. Determining a risk of osteoarthritis sometimes refers to determining whether an individual is at an increased risk of osteoarthritis (e.g., intermediate risk or higher risk).

[0083] Thus, featured herein is a method for identifying a subject who is at risk of osteoarthritis, which comprises detecting an aberration associated with osteoarthritis in a nucleic acid sample from the subject. An embodiment is a method for detecting a risk of osteoarthritis in a subject, which comprises detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; and (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic site; whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject. In certain embodiments, polymorphic variants at the positions described herein are detected for determining a risk of osteoarthritis, and polymorphic variants at positions in linkage disequilibrium with these positions are detected for determining a risk of osteoarthritis. As used herein, the terms "SEQ ID NO: 1-3" and other nucleotide sequences "referenced in Table A" refers to individual sequences in SEQ ID NO: 1, 2 or 3 or any individual sequence referenced in Table A, each sequence being separately applicable to embodiments described herein.

[0084] Risk of osteoarthritis sometimes is expressed as a probability, such as an odds ratio, percentage, or risk factor. Risk often is based upon the presence or absence of one or more polymorphic variants described herein, and also may be based in part upon phenotypic traits of the individual being tested. Methods for calculating risk based upon patient data are well known (*see, e.g., Agresti, Categorical Data Analysis*, 2nd Ed. 2002. Wiley). Allelotyping and genotyping analyses may be carried out in populations other than those exemplified herein to enhance the predictive power of the prognostic method. These further analyses are executed in view of the exemplified procedures described herein, and may be based upon the same polymorphic variations or additional polymorphic variations.

[0085] In certain embodiments, determining the presence of a combination of two or more polymorphic variants associated with osteoarthritis in one or more genetic loci (e.g., one or more genes) of the sample is determined to identify, quantify and/or estimate, risk of osteoarthritis. The risk often is the probability of having or developing osteoarthritis. The risk sometimes is expressed as a relative risk with respect to a population average risk of osteoarthritis, and sometimes is expressed as a relative risk with respect to the lowest risk group. Such relative risk assessments often are based upon penetrance values determined by statistical methods, and are particularly useful to clinicians and insurance companies for assessing risk of osteoarthritis (e.g., a clinician can target appropriate detection, prevention and therapeutic regimens to a patient after determining the patient's risk of osteoarthritis, and an insurance company can fine tune actuarial tables based upon population genotype assessments of osteoarthritis risk). Risk of osteoarthritis sometimes is expressed as an odds ratio, which is the odds of a particular person having a genotype has or will develop osteoarthritis with respect to another genotype group (e.g., the most disease protective genotype or population average). In related embodiments, the determination is utilized to identify a subject at risk of osteoarthritis. In an embodiment, two or more polymorphic variations are detected in two or more regions in human genomic DNA associated with increased risk of osteoarthritis, such as a locus containing a *KIAA0296*, *chrom 4* or *SNW1* or other locus referenced in Table A, for example. In certain embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected in the sample. In specific embodiments, polymorphic variants are detected in a *KIAA0296*, *chrom 4* or *SNW1* region or other region referenced in Table A, for example. In another embodiment, polymorphic variants are detected at two or three positions in a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A. In certain embodiments, polymorphic variants are detected at other genetic loci (e.g., the polymorphic variants can be detected in a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A in addition to other loci or only in other loci), where the other loci include but are not limited to those described in concurrently-filed patent applications having attorney docket number 524593008700, 524593008800, 524593008900, 524593009000, 524593009100 or 524593009200, each of which is incorporated herein by reference in its entirety.

[0086] Results from prognostic tests may be combined with other test results to diagnose osteoarthritis. For example, prognostic results may be gathered, a patient sample may be ordered based on a determined predisposition to osteoarthritis, the patient sample is analyzed, and the results of the analysis may be utilized to diagnose osteoarthritis. Also osteoarthritis diagnostic method can be developed from studies used to generate prognostic methods in which populations are stratified into subpopulations having different progressions of osteoarthritis. In another embodiment, prognostic results may be gathered, a patient's risk factors for developing osteoarthritis (e.g., age, weight, race, diet) analyzed, and a patient sample may be ordered based on a determined predisposition to osteoarthritis.

[0087] The nucleic acid sample typically is isolated from a biological sample obtained from a subject. For example, nucleic acid can be isolated from blood, saliva, sputum, urine, cell scrapings, and biopsy tissue. The nucleic acid sample can be isolated from a biological sample using standard techniques, such as the technique described in Example 2. As used herein, the term "subject" refers primarily to humans but also refers to other mammals such as dogs, cats, and ungulates (*e.g.*, cattle, sheep, and swine). Subjects also include avians (*e.g.*, chickens and turkeys), reptiles, and fish (*e.g.*, salmon), as embodiments described herein can be adapted to nucleic acid samples isolated from any of these organisms. The nucleic acid sample may be isolated from the subject and then directly utilized in a method for determining the presence of a polymorphic variant, or alternatively, the sample may be isolated and then stored (*e.g.*, frozen) for a period of time before being subjected to analysis.

[0088] The presence or absence of a polymorphic variant is determined using one or both chromosomal complements represented in the nucleic acid sample. Determining the presence or absence of a polymorphic variant in both chromosomal complements represented in a nucleic acid sample from a subject having a copy of each chromosome is useful for determining the zygosity of an individual for the polymorphic variant (*i.e.*, whether the individual is homozygous or heterozygous for the polymorphic variant). Any oligonucleotide-based diagnostic may be utilized to determine whether a sample includes the presence or absence of a polymorphic variant in a sample. For example, primer extension methods, ligase sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,679,524 and 5,952,174, and WO 01/27326), mismatch sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,851,770; 5,958,692; 6,110,684; and 6,183,958), microarray sequence determination methods, restriction fragment length polymorphism (RFLP), single strand conformation polymorphism detection (SSCP) (*e.g.*, U.S. Pat. Nos. 5,891,625 and 6,013,499), PCR-based assays (*e.g.*, TAQMAN<sup>®</sup> PCR System (Applied Biosystems)), and nucleotide sequencing methods may be used.

[0089] Oligonucleotide extension methods typically involve providing a pair of oligonucleotide primers in a polymerase chain reaction (PCR) or in other nucleic acid amplification methods for the purpose of amplifying a region from the nucleic acid sample that comprises the polymorphic variation. One oligonucleotide primer is complementary to a region 3' of the polymorphism and the other is complementary to a region 5' of the polymorphism. A PCR primer pair may be used in methods disclosed in U.S. Pat. Nos. 4,683,195; 4,683,202; 4,965,188; 5,656,493; 5,998,143; 6,140,054; WO 01/27327; and WO 01/27329 for example. PCR primer pairs may also be used in any commercially available machines that perform PCR, such as any of the GENEAMP<sup>®</sup> Systems available from Applied Biosystems. Also, those of ordinary skill in the art will be able to design oligonucleotide primers based upon a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A using knowledge available in the art.

[0090] Also provided is an extension oligonucleotide that hybridizes to the amplified fragment adjacent to the polymorphic variation. As used herein, the term "adjacent" refers to the 3' end of the extension oligonucleotide being often 1 nucleotide from the 5' end of the polymorphic site, and sometimes 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the polymorphic site, in the nucleic acid when the extension oligonucleotide is hybridized to the nucleic acid. The extension oligonucleotide then is extended by one or more nucleotides, and the number and/or type of nucleotides that are added to the extension oligonucleotide determine whether the polymorphic variant is present. Oligonucleotide extension methods are disclosed, for example, in U.S. Pat. Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 6,013,431; 6,017,702; 6,046,005; 6,087,095; 6,210,891; and WO 01/20039. Oligonucleotide extension methods using mass spectrometry are described, for example, in U.S. Pat. Nos. 5,547,835; 5,605,798; 5,691,141; 5,849,542; 5,869,242; 5,928,906; 6,043,031; and 6,194,144, and a method often utilized is described herein in Example 2.

[0091] A microarray can be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A microarray may include any oligonucleotides described herein, and methods for making and using oligonucleotide microarrays suitable for diagnostic use are disclosed in U.S. Pat. Nos. 5,492,806; 5,525,464; 5,589,330; 5,695,940; 5,849,483; 6,018,041; 6,045,996; 6,136,541; 6,142,681; 6,156,501; 6,197,506; 6,223,127; 6,225,625; 6,229,911; 6,239,273; WO 00/52625; WO 01/25485; and WO 01/29259. The microarray typically comprises a solid support and the oligonucleotides may be linked to this solid support by covalent bonds or by non-covalent interactions. The oligonucleotides may also be linked to the solid support directly or by a spacer molecule. A microarray may comprise one or more oligonucleotides complementary to a polymorphic site set forth herein.

[0092] A kit also may be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A kit often comprises one or more pairs of oligonucleotide primers useful for amplifying a fragment of a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A or a substantially identical sequence thereof, where the fragment includes a polymorphic site. The kit sometimes comprises a polymerizing agent, for example, a thermostable nucleic acid polymerase such as one disclosed in U.S. Pat. Nos. 4,889,818 or 6,077,664. Also, the kit often comprises an elongation oligonucleotide that hybridizes to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A in a nucleic acid sample adjacent to the polymorphic site. Where the kit includes an elongation oligonucleotide, it also often comprises chain elongating nucleotides, such as dATP, dTTP, dGTP, dCTP, and dITP, including analogs of dATP, dTTP, dGTP, dCTP and dITP, provided that such analogs are substrates for a thermostable nucleic acid polymerase and can be incorporated into a nucleic acid chain elongated from the extension oligonucleotide. Along with chain



elongating nucleotides would be one or more chain terminating nucleotides such as ddATP, ddTTP, ddGTP, ddCTP, and the like. In an embodiment, the kit comprises one or more oligonucleotide primer pairs, a polymerizing agent, chain elongating nucleotides, at least one elongation oligonucleotide, and one or more chain terminating nucleotides. Kits optionally include buffers, vials, microtiter plates, and instructions for use.

[0093] An individual identified as being at risk of osteoarthritis may be heterozygous or homozygous with respect to the allele associated with a higher risk of osteoarthritis. A subject homozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively high risk of osteoarthritis, a subject heterozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively intermediate risk of osteoarthritis, and a subject homozygous for an allele associated with a decreased risk of osteoarthritis is at a comparatively low risk of osteoarthritis. A genotype may be assessed for a complementary strand, such that the complementary nucleotide at a particular position is detected.

[0094] Also featured are methods for determining risk of osteoarthritis and/or identifying a subject at risk of osteoarthritis by contacting a polypeptide or protein encoded by a *KLAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A from a subject with an antibody that specifically binds to an epitope associated with increased risk of osteoarthritis in the polypeptide.

#### Applications of Prognostic and Diagnostic Results to Pharmacogenomic Methods

[0095] Pharmacogenomics is a discipline that involves tailoring a treatment for a subject according to the subject's genotype as a particular treatment regimen may exert a differential effect depending upon the subject's genotype. For example, based upon the outcome of a prognostic test described herein, a clinician or physician may target pertinent information and preventative or therapeutic treatments to a subject who would be benefited by the information or treatment and avoid directing such information and treatments to a subject who would not be benefited (*e.g.*, the treatment has no therapeutic effect and/or the subject experiences adverse side effects).

[0096] The following is an example of a pharmacogenomic embodiment. A particular treatment regimen can exert a differential effect depending upon the subject's genotype. Where a candidate therapeutic exhibits a significant interaction with a major allele and a comparatively weak interaction with a minor allele (*e.g.*, an order of magnitude or greater difference in the interaction), such a therapeutic typically would not be administered to a subject genotyped as being homozygous for the minor allele, and sometimes not administered to a subject genotyped as being heterozygous for the minor allele. In another example, where a candidate therapeutic is not significantly toxic when administered to subjects who are homozygous for a major allele but is comparatively toxic when administered to subjects

heterozygous or homozygous for a minor allele, the candidate therapeutic is not typically administered to subjects who are genotyped as being heterozygous or homozygous with respect to the minor allele.

[0097] The methods described herein are applicable to pharmacogenomic methods for preventing, alleviating or treating osteoarthritis. For example, a nucleic acid sample from an individual may be subjected to a prognostic test described herein. Where one or more polymorphic variations associated with increased risk of osteoarthritis are identified in a subject, information for preventing or treating osteoarthritis and/or one or more osteoarthritis treatment regimens then may be prescribed to that subject.

[0098] In certain embodiments, a treatment or preventative regimen is specifically prescribed and/or administered to individuals who will most benefit from it based upon their risk of developing osteoarthritis assessed by the methods described herein. Thus, provided are methods for identifying a subject predisposed to osteoarthritis and then prescribing a therapeutic or preventative regimen to individuals identified as having a predisposition. Thus, certain embodiments are directed to a method for reducing osteoarthritis in a subject, which comprises: detecting the presence or absence of a polymorphic variant associated with osteoarthritis in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; and (d) a fragment of a polynucleotide sequence of (a), (b), or (c); and prescribing or administering a treatment regimen to a subject from whom the sample originated where the presence of a polymorphic variation associated with osteoarthritis is detected in the nucleotide sequence. In these methods, predisposition results may be utilized in combination with other test results to diagnose osteoarthritis.

[0099] Certain preventative treatments often are prescribed to subjects having a predisposition to osteoarthritis and where the subject is diagnosed with osteoarthritis or is diagnosed as having symptoms indicative of an early stage of osteoarthritis. The treatment sometimes is preventative (e.g., is prescribed or administered to reduce the probability that osteoarthritis arises or progresses), sometimes is therapeutic, and sometimes delays, alleviates or halts the progression of osteoarthritis. Any known preventative or therapeutic treatment for alleviating or preventing the occurrence of osteoarthritis is prescribed and/or administered. For example, the treatment often is directed to decreasing pain and improving joint movement. Examples of OA treatments include exercises to keep joints flexible and improve muscle strength. Different medications to control pain, including corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs, e.g., Voltaren); cyclooxygenase-2 (COX-2) inhibitors

(e.g., Celebrex, Vioxx, Mobic, and Bextra); monoclonal antibodies (e.g., Remicade); tumor necrosis factor inhibitors (e.g., Enbrel); or injections of glucocorticoids, hyaluronic acid or chondroitin sulfate into joints that are inflamed and not responsive to NSAIDS. Orally administered chondroitin sulfate also may be used as a therapeutic, as it may increase hyaluronic acid levels and viscosity of synovial fluid, and decrease collagenase levels in synovial fluid. Also, glucosamine can serve as an OA therapeutic as delivering it into joints may inhibit enzymes involved in cartilage degradation and enhance the production of hyaluronic acid. For mild pain without inflammation, acetaminophen may be used. Other treatments include: heat/cold therapy for temporary pain relief; joint protection to prevent strain or stress on painful joints; surgery to relieve chronic pain in damaged joints; and weight control to prevent extra stress on weight-bearing joints.

[0100] As therapeutic approaches for treating osteoarthritis continue to evolve and improve, the goal of treatments for osteoarthritis related disorders is to intervene even before clinical signs first manifest. Thus, genetic markers associated with susceptibility to osteoarthritis prove useful for early diagnosis, prevention and treatment of osteoarthritis.

[0101] As osteoarthritis preventative and treatment information can be specifically targeted to subjects in need thereof (e.g., those at risk of developing osteoarthritis or those in an early stage of osteoarthritis), provided herein is a method for preventing or reducing the risk of developing osteoarthritis in a subject, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying a subject with a predisposition to osteoarthritis, whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject; and (c) if such a predisposition is identified, providing the subject with information about methods or products to prevent or reduce osteoarthritis or to delay the onset of osteoarthritis. Also provided is a method of targeting information or advertising to a subpopulation of a human population based on the subpopulation being genetically predisposed to a disease or condition, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; and (c) providing information only to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition.

[0102] Pharmacogenomics methods also may be used to analyze and predict a response to osteoarthritis treatment or a drug. For example, if pharmacogenomics analysis indicates a likelihood that an individual will respond positively to osteoarthritis treatment with a particular drug, the drug may be administered to the individual. Conversely, if the analysis indicates that an individual is likely to respond negatively to treatment with a particular drug, an alternative course of treatment may be prescribed. A

negative response may be defined as either the absence of an efficacious response or the presence of toxic side effects. The response to a therapeutic treatment can be predicted in a background study in which subjects in any of the following populations are genotyped: a population that responds favorably to a treatment regimen, a population that does not respond significantly to a treatment regimen, and a population that responds adversely to a treatment regimen (*e.g.*, exhibits one or more side effects). These populations are provided as examples and other populations and subpopulations may be analyzed. Based upon the results of these analyses, a subject is genotyped to predict whether he or she will respond favorably to a treatment regimen, not respond significantly to a treatment regimen, or respond adversely to a treatment regimen.

[0103] The tests described herein also are applicable to clinical drug trials. One or more polymorphic variants indicative of response to an agent for treating osteoarthritis or to side effects to an agent for treating osteoarthritis may be identified using the methods described herein. Thereafter, potential participants in clinical trials of such an agent may be screened to identify those individuals most likely to respond favorably to the drug and exclude those likely to experience side effects. In that way, the effectiveness of drug treatment may be measured in individuals who respond positively to the drug, without lowering the measurement as a result of the inclusion of individuals who are unlikely to respond positively in the study and without risking undesirable safety problems.

[0104] Thus, another embodiment is a method of selecting an individual for inclusion in a clinical trial of a treatment or drug comprising the steps of: (a) obtaining a nucleic acid sample from an individual; (b) determining the identity of a polymorphic variation which is associated with a positive response to the treatment or the drug, or at least one polymorphic variation which is associated with a negative response to the treatment or the drug in the nucleic acid sample, and (c) including the individual in the clinical trial if the nucleic acid sample contains said polymorphic variation associated with a positive response to the treatment or the drug or if the nucleic acid sample lacks said polymorphic variation associated with a negative response to the treatment or the drug. In addition, the methods described herein for selecting an individual for inclusion in a clinical trial of a treatment or drug encompass methods with any further limitation described in this disclosure, or those following, specified alone or in any combination. The polymorphic variation may be in a sequence selected individually or in any combination from the group consisting of (i) a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (ii) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; and (iv) a fragment of a polynucleotide sequence of (i), (ii), or (iii) comprising the polymorphic site. The including

step (c) optionally comprises administering the drug or the treatment to the individual if the nucleic acid sample contains the polymorphic variation associated with a positive response to the treatment or the drug and the nucleic acid sample lacks said biallelic marker associated with a negative response to the treatment or the drug.

[0105] Also provided herein is a method of partnering between a diagnostic/prognostic testing provider and a provider of a consumable product, which comprises: (a) the diagnostic/prognostic testing provider detects the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) the diagnostic/prognostic testing provider identifies the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; (c) the diagnostic/prognostic testing provider forwards information to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition; and (d) the provider of a consumable product forwards to the diagnostic test provider a fee every time the diagnostic/prognostic test provider forwards information to the subject as set forth in step (c) above.

#### Compositions Comprising Osteoarthritis-Directed Molecules

[0106] Featured herein is a composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and one or more molecules specifically directed and targeted to a nucleic acid comprising a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence, other nucleotide sequence referenced in Table A, or an encoded amino acid sequence. Such directed molecules include, but are not limited to, a compound that binds to a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence, or other nucleotide sequence referenced in Table A, or encoded amino acid sequence; a RNAi or siRNA molecule having a strand complementary or substantially complementary to a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A (e.g., hybridizes to a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A under conditions of high stringency); an antisense nucleic acid complementary or substantially complementary to an RNA encoded by a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A (e.g., hybridizes to a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A under conditions of high stringency); a ribozyme that hybridizes to a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A (e.g., hybridizes to a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A under conditions of high stringency); a nucleic acid aptamer that specifically binds a polypeptide encoded by a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A; and an antibody that specifically binds to a polypeptide encoded by a *KIAA0296, chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A or

binds to a nucleic acid having such a nucleotide sequence. In an embodiment, the antibody selectively binds to an epitope comprising an amino acid encoded by rs734784, rs1042164, rs749670, rs955592 and rs1040461. In specific embodiments, the osteoarthritis directed molecule interacts with a nucleic acid or polypeptide variant associated with osteoarthritis, such as variants referenced herein. In other embodiments, the osteoarthritis directed molecule interacts with a polypeptide involved in a signal pathway of a polypeptide encoded by a *KLAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a nucleic acid comprising such a nucleotide sequence.

[0107] Compositions sometimes include an adjuvant known to stimulate an immune response, and in certain embodiments, an adjuvant that stimulates a T-cell lymphocyte response. Adjuvants are known, including but not limited to an aluminum adjuvant (e.g., aluminum hydroxide); a cytokine adjuvant or adjuvant that stimulates a cytokine response (e.g., interleukin (IL)-12 and/or gamma-interferon cytokines); a Freund-type mineral oil adjuvant emulsion (e.g., Freund's complete or incomplete adjuvant); a synthetic lipid compound; a copolymer adjuvant (e.g., TitreMax); a saponin; Quil A; a liposome; an oil-in-water emulsion (e.g., an emulsion stabilized by Tween 80 and pluronic polyoxyethylene/polyoxypropylene block copolymer (Syntex Adjuvant Formulation); TitreMax; detoxified endotoxin (MPL) and mycobacterial cell wall components (TDW, CWS) in 2% squalene (Ribi Adjuvant System)); a muramyl dipeptide; an immune-stimulating complex (ISCOM, e.g., an Ag-modified saponin/cholesterol micelle that forms stable cage-like structure); an aqueous phase adjuvant that does not have a depot effect (e.g., Gerbu adjuvant); a carbohydrate polymer (e.g., AdjuPrime); L-tyrosine; a manide-oleate compound (e.g., Montanide); an ethylene-vinyl acetate copolymer (e.g., Elvax 40W1,2); or lipid A, for example. Such compositions are useful for generating an immune response against osteoarthritis directed molecule (e.g., an HLA-binding subsequence within a polypeptide encoded by a *KLAA0296*, *chrom 4* or *SNW1* nucleotide sequence). In such methods, a peptide having an amino acid subsequence of a polypeptide encoded by a *KLAA0296*, *chrom 4* or *SNW1* nucleotide sequence is delivered to a subject, where the subsequence binds to an HLA molecule and induces a CTL lymphocyte response. The peptide sometimes is delivered to the subject as an isolated peptide or as a minigene in a plasmid that encodes the peptide. Methods for identifying HLA-binding subsequences in such polypeptides are known (see e.g., publication WO02/20616 and PCT application US98/01373 for methods of identifying such sequences).

[0108] The cell may be in a group of cells cultured *in vitro* or in a tissue maintained *in vitro* or present in an animal *in vivo* (e.g., a rat, mouse, ape or human). In certain embodiments, a composition comprises a component from a cell such as a nucleic acid molecule (e.g., genomic DNA), a protein mixture or isolated protein, for example. The aforementioned compositions have utility in diagnostic, prognostic and pharmacogenomic methods described previously and in therapeutics described hereafter. Certain osteoarthritis directed molecules are described in greater detail below.

### Compounds

[0109] Compounds can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive (see, e.g., Zuckermann et al., J. Med. Chem. 37: 2678-85 (1994)); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; "one-bead one-compound" library methods; and synthetic library methods using affinity chromatography selection. Biological library and peptoid library approaches are typically limited to peptide libraries, while the other approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, Anticancer Drug Des. 12: 145, (1997)). Examples of methods for synthesizing molecular libraries are described, for example, in DeWitt et al., Proc. Natl. Acad. Sci. U.S.A. 90: 6909 (1993); Erb et al., Proc. Natl. Acad. Sci. USA 91: 11422 (1994); Zuckermann et al., J. Med. Chem. 37: 2678 (1994); Cho et al., Science 261: 1303 (1993); Carrell et al., Angew. Chem. Int. Ed. Engl. 33: 2059 (1994); Carell et al., Angew. Chem. Int. Ed. Engl. 33: 2061 (1994); and in Gallop et al., J. Med. Chem. 37: 1233 (1994).

[0110] Libraries of compounds may be presented in solution (e.g., Houghten, Biotechniques 13: 412-421 (1992)), or on beads (Lam, Nature 354: 82-84 (1991)), chips (Fodor, Nature 364: 555-556 (1993)), bacteria or spores (Ladner, United States Patent No. 5,223,409), plasmids (Cull et al., Proc. Natl. Acad. Sci. USA 89: 1865-1869 (1992)) or on phage (Scott and Smith, Science 249: 386-390 (1990); Devlin, Science 249: 404-406 (1990); Cwirla et al., Proc. Natl. Acad. Sci. 87: 6378-6382 (1990); Felici, J. Mol. Biol. 222: 301-310 (1991); Ladner supra.).

[0111] A compound sometimes alters expression and sometimes alters activity of a polypeptide target and may be a small molecule. Small molecules include, but are not limited to, peptides, peptidomimetics (e.g., peptoids), amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

### Antisense Nucleic Acid Molecules, Ribozymes, RNAi, siRNA and Modified Nucleic Acid Molecules

[0112] An "antisense" nucleic acid refers to a nucleotide sequence complementary to a "sense" nucleic acid encoding a polypeptide, e.g., complementary to the coding strand of a double-stranded

cDNA molecule or complementary to an mRNA sequence. The antisense nucleic acid can be complementary to an entire coding strand, or to a portion thereof or a substantially identical sequence thereof. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence (e.g., 5' and 3' untranslated regions in SEQ ID NO: 1-3 or a nucleotide sequence referenced in Table A).

[0113] An antisense nucleic acid can be designed such that it is complementary to the entire coding region of an mRNA encoded by a nucleotide sequence (e.g., SEQ ID NO: 1-3 or a nucleotide sequence referenced in Table A), and often the antisense nucleic acid is an oligonucleotide antisense to only a portion of a coding or noncoding region of the mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of the mRNA, e.g., between the -10 and +10 regions of the target gene nucleotide sequence of interest. An antisense oligonucleotide can be, for example, about 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more nucleotides in length. The antisense nucleic acids, which include the ribozymes described hereafter, can be designed to target a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence, often a variant associated with osteoarthritis, or a substantially identical sequence thereof. Among the variants, minor alleles and major alleles can be targeted, and those associated with a higher risk of osteoarthritis are often designed, tested, and administered to subjects.

[0114] An antisense nucleic acid can be constructed using chemical synthesis and enzymatic ligation reactions using standard procedures. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Antisense nucleic acid also can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

[0115] When utilized as therapeutics, antisense nucleic acids typically are administered to a subject (e.g., by direct injection at a tissue site) or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a polypeptide and thereby inhibit expression of the polypeptide, for example, by inhibiting transcription and/or translation. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then are administered systemically. For systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, for example, by linking antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. Antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. Sufficient intracellular



concentrations of antisense molecules are achieved by incorporating a strong promoter, such as a pol II or pol III promoter, in the vector construct.

[0116] Antisense nucleic acid molecules sometimes are alpha-anomeric nucleic acid molecules. An alpha-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual beta-units, the strands run parallel to each other (Gaultier et al., Nucleic Acids. Res. 15: 6625-6641 (1987)). Antisense nucleic acid molecules can also comprise a 2'-o-methylribonucleotide (Inoue et al., Nucleic Acids Res. 15: 6131-6148 (1987)) or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett. 215: 327-330 (1987)). Antisense nucleic acids sometimes are composed of DNA or PNA or any other nucleic acid derivatives described previously.

[0117] In another embodiment, an antisense nucleic acid is a ribozyme. A ribozyme having specificity for a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A can include one or more sequences complementary to such a nucleotide sequence, and a sequence having a known catalytic region responsible for mRNA cleavage (see e.g., U.S. Pat. No. 5,093,246 or Haselhoff and Gerlach, Nature 334: 585-591 (1988)). For example, a derivative of a Tetrahymena L-19 IVS RNA is sometimes utilized in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a mRNA (see e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742). Also, target mRNA sequences can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see e.g., Bartel & Szostak, Science 261: 1411-1418 (1993)).

[0118] Osteoarthritis directed molecules include in certain embodiments nucleic acids that can form triple helix structures with a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical sequence thereof, especially one that includes a regulatory region that controls expression of a polypeptide. Gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of a nucleotide sequence referenced herein or a substantially identical sequence (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of a gene in target cells (see e.g., Helene, Anticancer Drug Des. 6(6): 569-84 (1991); Helene et al., Ann. N.Y. Acad. Sci. 660: 27-36 (1992); and Maher, Bioassays 14(12): 807-15 (1992). Potential sequences that can be targeted for triple helix formation can be increased by creating a so-called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

[0119] Osteoarthritis directed molecules include RNAi and siRNA nucleic acids. Gene expression may be inhibited by the introduction of double-stranded RNA (dsRNA), which induces potent and specific gene silencing, a phenomenon called RNA interference or RNAi. See, e.g., Fire et al., US Patent

Number 6,506,559; Tuschl et al. PCT International Publication No. WO 01/75164; Kay et al. PCT International Publication No. WO 03/010180A1; or Bosher JM, Labouesse, Nat Cell Biol 2000 Feb;2(2):E31-6. This process has been improved by decreasing the size of the double-stranded RNA to 20-24 base pairs (to create small-interfering RNAs or siRNAs) that "switched off" genes in mammalian cells without initiating an acute phase response, i.e., a host defense mechanism that often results in cell death (see, e.g., Caplen et al. Proc Natl Acad Sci U S A. 2001 Aug 14;98(17):9742-7 and Elbashir et al. Methods 2002 Feb;26(2):199-213). There is increasing evidence of post-transcriptional gene silencing by RNA interference (RNAi) for inhibiting targeted expression in mammalian cells at the mRNA level, in human cells. There is additional evidence of effective methods for inhibiting the proliferation and migration of tumor cells in human patients, and for inhibiting metastatic cancer development (see, e.g., U.S. Patent Application No. US2001000993183; Caplen et al. Proc Natl Acad Sci U S A; and Abderrahmani et al. Mol Cell Biol 2001 Nov21(21):7256-67).

**[0120]** An "siRNA" or "RNAi" refers to a nucleic acid that forms a double stranded RNA and has the ability to reduce or inhibit expression of a gene or target gene when the siRNA is delivered to or expressed in the same cell as the gene or target gene. "siRNA" refers to short double-stranded RNA formed by the complementary strands. Complementary portions of the siRNA that hybridize to form the double stranded molecule often have substantial or complete identity to the target molecule sequence. In one embodiment, an siRNA refers to a nucleic acid that has substantial or complete identity to a target gene and forms a double stranded siRNA.

**[0121]** When designing the siRNA molecules, the targeted region often is selected from a given DNA sequence beginning 50 to 100 nucleotides downstream of the start codon. See, e.g., Elbashir et al., Methods 26:199-213 (2002). Initially, 5' or 3' UTRs and regions nearby the start codon were avoided assuming that UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP or RISC endonuclease complex. Sometimes regions of the target 23 nucleotides in length conforming to the sequence motif AA(N19)TT (N, an nucleotide), and regions with approximately 30% to 70% G/C-content (often about 50% G/C-content) often are selected. If no suitable sequences are found, the search often is extended using the motif NA(N21). The sequence of the sense siRNA sometimes corresponds to (N19) TT or N21 (position 3 to 23 of the 23-nt motif), respectively. In the latter case, the 3' end of the sense siRNA often is converted to TT. The rationale for this sequence conversion is to generate a symmetric duplex with respect to the sequence composition of the sense and antisense 3' overhangs. The antisense siRNA is synthesized as the complement to position 1 to 21 of the 23-nt motif. Because position 1 of the 23-nt motif is not recognized sequence-specifically by the antisense siRNA, the 3'-most nucleotide residue of the antisense siRNA can be chosen deliberately. However, the penultimate nucleotide of the antisense siRNA (complementary to position 2 of the 23-nt motif) often is complementary to the targeted sequence. For simplifying chemical synthesis, TT often is

utilized. siRNAs corresponding to the target motif NAR(N17)YNN, where R is purine (A,G) and Y is pyrimidine (C,U), often are selected. Respective 21 nucleotide sense and antisense siRNAs often begin with a purine nucleotide and can also be expressed from pol III expression vectors without a change in targeting site. Expression of RNAs from pol III promoters often is efficient when the first transcribed nucleotide is a purine.

[0122] The sequence of the siRNA can correspond to the full length target gene, or a subsequence thereof. Often, the siRNA is about 15 to about 50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, sometimes about 20-30 nucleotides in length or about 20-25 nucleotides in length, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length. The siRNA sometimes is about 21 nucleotides in length. Methods of using siRNA are well known in the art, and specific siRNA molecules may be purchased from a number of companies including Dharmacon Research, Inc.

[0123] Antisense, ribozyme, RNAi and siRNA nucleic acids can be altered to form modified nucleic acid molecules. The nucleic acids can be altered at base moieties, sugar moieties or phosphate backbone moieties to improve stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup et al., *Bioorganic & Medicinal Chemistry* 4 (1): 5-23 (1996)). As used herein, the terms "peptide nucleic acid" or "PNA" refers to a nucleic acid mimic such as a DNA mimic, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of a PNA can allow for specific hybridization to DNA and RNA under conditions of low ionic strength. Synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described, for example, in Hyrup et al., (1996) supra and Perry-O'Keefe et al., *Proc. Natl. Acad. Sci.* 93: 14670-675 (1996).

[0124] PNA nucleic acids can be used in prognostic, diagnostic, and therapeutic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNA nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as "artificial restriction enzymes" when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup (1996) supra)); or as probes or primers for DNA sequencing or hybridization (Hyrup et al., (1996) supra; Perry-O'Keefe supra).

[0125] In other embodiments, oligonucleotides may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across cell membranes (see e.g., Letsinger et al., *Proc. Natl. Acad. Sci. USA* 86: 6553-6556 (1989); Lemaître et al., *Proc. Natl. Acad. Sci. USA* 84: 648-652 (1987); PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-

triggered cleavage agents (See, e.g., Krol et al., Bio-Techniques 6: 958-976 (1988)) or intercalating agents. (See, e.g., Zon, Pharm. Res. 5: 539-549 (1988) ). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

[0126] Also included herein are molecular beacon oligonucleotide primer and probe molecules having one or more regions complementary to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical sequence thereof, two complementary regions one having a fluorophore and one a quencher such that the molecular beacon is useful for quantifying the presence of the nucleic acid in a sample. Molecular beacon nucleic acids are described, for example, in Lizardi et al., U.S. Patent No. 5,854,033; Nazarenko et al., U.S. Patent No. 5,866,336, and Livak et al., U.S. Patent 5,876,930.

#### Antibodies

[0127] The term "antibody" as used herein refers to an immunoglobulin molecule or immunologically active portion thereof, i.e., an antigen-binding portion. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')<sub>2</sub> fragments which can be generated by treating the antibody with an enzyme such as pepsin. An antibody sometimes is a polyclonal, monoclonal, recombinant (e.g., a chimeric or humanized), fully human, non-human (e.g., murine), or a single chain antibody. An antibody may have effector function and can fix complement, and is sometimes coupled to a toxin or imaging agent.

[0128] A full-length polypeptide or antigenic peptide fragment encoded by a nucleotide sequence referenced herein can be used as an immunogen or can be used to identify antibodies made with other immunogens, e.g., cells, membrane preparations, and the like. An antigenic peptide often includes at least 8 amino acid residues of the amino acid sequences encoded by a nucleotide sequence referenced herein, or substantially identical sequence thereof, and encompasses an epitope. Antigenic peptides sometimes include 10 or more amino acids, 15 or more amino acids, 20 or more amino acids, or 30 or more amino acids. Hydrophilic and hydrophobic fragments of polypeptides sometimes are used as immunogens.

[0129] Epitopes encompassed by the antigenic peptide are regions located on the surface of the polypeptide (e.g., hydrophilic regions) as well as regions with high antigenicity. For example, an Emini surface probability analysis of the human polypeptide sequence can be used to indicate the regions that have a particularly high probability of being localized to the surface of the polypeptide and are thus likely to constitute surface residues useful for targeting antibody production. The antibody may bind an epitope on any domain or region on polypeptides described herein.

[0130] Also, chimeric, humanized, and completely human antibodies are useful for applications which include repeated administration to subjects. Chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, can be made using standard recombinant DNA techniques. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson et al International Application No. PCT/US86/02269; Akira, et al European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al European Patent Application 173,494; Neuberger et al PCT International Publication No. WO 86/01533; Cabilly et al U.S. Patent No. 4,816,567; Cabilly et al European Patent Application 125,023; Better et al., Science 240: 1041-1043 (1988); Liu et al., Proc. Natl. Acad. Sci. USA 84: 3439-3443 (1987); Liu et al., J. Immunol. 139: 3521-3526 (1987); Sun et al., Proc. Natl. Acad. Sci. USA 84: 214-218 (1987); Nishimura et al., Canc. Res. 47: 999-1005 (1987); Wood et al., Nature 314: 446-449 (1985); and Shaw et al., J. Natl. Cancer Inst. 80: 1553-1559 (1988); Morrison, S. L., Science 229: 1202-1207 (1985); Oi et al., BioTechniques 4: 214 (1986); Winter U.S. Patent 5,225,539; Jones et al., Nature 321: 552-525 (1986); Verhoeyan et al., Science 239: 1534; and Beidler et al., J. Immunol. 141: 4053-4060 (1988).

[0131] Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice that are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. See, for example, Lonberg and Huszar, Int. Rev. Immunol. 13: 65-93 (1995); and U.S. Patent Nos. 5,625,126; 5,633,425; 5,569,825; 5,661,016; and 5,545,806. In addition, companies such as Abgenix, Inc. (Fremont, CA) and Medarex, Inc. (Princeton, NJ), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above. Completely human antibodies that recognize a selected epitope also can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody (e.g., a murine antibody) is used to guide the selection of a completely human antibody recognizing the same epitope. This technology is described for example by Jaspers et al., Bio/Technology 12: 899-903 (1994).

[0132] An antibody can be a single chain antibody. A single chain antibody (scFv) can be engineered (see, e.g., Colcher et al., Ann. N Y Acad. Sci. 880: 263-80 (1999); and Reiter, Clin. Cancer Res. 2: 245-52 (1996)). Single chain antibodies can be dimerized or multimerized to generate multivalent antibodies having specificities for different epitopes of the same target polypeptide.

[0133] Antibodies also may be selected or modified so that they exhibit reduced or no ability to bind an Fc receptor. For example, an antibody may be an isotype or subtype, fragment or other mutant, which does not support binding to an Fc receptor (e.g., it has a mutagenized or deleted Fc receptor binding region).

[0134] Also, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1 dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thiotepa chlorambucil, melphalan, carmustine (BCNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

[0135] Antibody conjugates can be used for modifying a given biological response. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, gamma-interferon, alpha-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Also, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980, for example.

[0136] An antibody (e.g., monoclonal antibody) can be used to isolate target polypeptides by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, an antibody can be used to detect a target polypeptide (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the polypeptide. Antibodies can be used diagnostically to monitor polypeptide levels in tissue as part of a clinical testing procedure, e.g., to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance (i.e., antibody labeling). Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase,  $\beta$ -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol;

examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{35}\text{S}$  or  $^3\text{H}$ . Also, an antibody can be utilized as a test molecule for determining whether it can treat osteoarthritis, and as a therapeutic for administration to a subject for treating osteoarthritis.

[0137] An antibody can be made by immunizing with a purified antigen, or a fragment thereof, e.g., a fragment described herein, a membrane associated antigen, tissues, e.g., crude tissue preparations, whole cells, preferably living cells, lysed cells, or cell fractions.

[0138] Included herein are antibodies which bind only a native polypeptide, only denatured or otherwise non-native polypeptide, or which bind both, as well as those having linear or conformational epitopes. Conformational epitopes sometimes can be identified by selecting antibodies that bind to native but not denatured polypeptide. Also featured are antibodies that specifically bind to a polypeptide variant associated with osteoarthritis.

#### Methods for Identifying Candidate Therapeutics for Treating Osteoarthritis

[0139] Current therapies for the treatment of osteoarthritis have limited efficacy, limited tolerability and significant mechanism-based side effects, and few of the available therapies adequately address underlying defects. Current therapeutic approaches were largely developed in the absence of defined molecular targets or even a solid understanding of disease pathogenesis. Therefore, provided are methods of identifying candidate therapeutics that target biochemical pathways related to the development of osteoarthritis.

[0140] Thus, featured herein are methods for identifying a candidate therapeutic for treating osteoarthritis. The methods comprise contacting a test molecule with a target molecule in a system. A "target molecule" as used herein refers to a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleotide sequence referenced in Table A, a substantially identical nucleic acid thereof, or a fragment thereof, and an encoded polypeptide of the foregoing. The methods also comprise determining the presence or absence of an interaction between the test molecule and the target molecule, where the presence of an interaction between the test molecule and the nucleic acid or polypeptide identifies the test molecule as a candidate osteoarthritis therapeutic. The interaction between the test molecule and the target molecule may be quantified.

[0141] Test molecules and candidate therapeutics include, but are not limited to, compounds, antisense nucleic acids, siRNA molecules, ribozymes, polypeptides or proteins encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical sequence or fragment thereof, and immunotherapeutics (e.g., antibodies and HLA-presented polypeptide fragments). A test molecule or candidate therapeutic may act as a modulator of target molecule concentration or target molecule function in a system. A "modulator" may agonize (i.e.,

up-regulates) or antagonize (i.e., down-regulates) a target molecule concentration partially or completely in a system by affecting such cellular functions as DNA replication and/or DNA processing (e.g., DNA methylation or DNA repair), RNA transcription and/or RNA processing (e.g., removal of intronic sequences and/or translocation of spliced mRNA from the nucleus), polypeptide production (e.g., translation of the polypeptide from mRNA), and/or polypeptide post-translational modification (e.g., glycosylation, phosphorylation, and proteolysis of pro-polypeptides). A modulator may also agonize or antagonize a biological function of a target molecule partially or completely, where the function may include adopting a certain structural conformation, interacting with one or more binding partners, ligand binding, catalysis (e.g., phosphorylation, dephosphorylation, hydrolysis, methylation, and isomerization), and an effect upon a cellular event (e.g., effecting progression of osteoarthritis).

[0142] As used herein, the term "system" refers to a cell free *in vitro* environment and a cell-based environment such as a collection of cells, a tissue, an organ, or an organism. A system is "contacted" with a test molecule in a variety of manners, including adding molecules in solution and allowing them to interact with one another by diffusion, cell injection, and any administration routes in an animal. As used herein, the term "interaction" refers to an effect of a test molecule on test molecule, where the effect sometimes is binding between the test molecule and the target molecule, and sometimes is an observable change in cells, tissue, or organism.

[0143] There are many standard methods for detecting the presence or absence of interaction between a test molecule and a target molecule. For example, titrametric, acidimetric, radiometric, NMR, monolayer, polarographic, spectrophotometric, fluorescent, and ESR assays probative of a target molecule interaction may be utilized. Any modulator can be tested in such methods and modulators for certain targets described in Table A are known. For example, modulators of protein tyrosine phosphatases (e.g., *PTPNI* includes a protein phosphatase domain) are described in WO-03072537, WO-03020688, WO-00218321, WO-00218323, WO-03055883, WO-03041729, WO-00226707, WO-00226743 and WO-03037328; modulators of potassium channels (e.g., *KCNSI* includes a potassium channel domain) are described in WO-09962891, WO-09716437, WO-09521813, WO-09521823, WO-09521824, WO-09521825 and WO-03088908; modulators of annexin (e.g., *ANXA6* includes an annexin domain) are described in WO-2004018670, WO-02067857, WO-2004013303 and WO-00147510; and modulators of protein kinases (e.g., *FYN* is a protein kinase) are described in WO-03081210, WO-02080926, WO-02076986, WO-03077921, WO03026666, WO03026665 and WO03026664.

[0144] Test molecule/target molecule interactions can be detected and/or quantified using assays known in the art. For example, an interaction can be determined by labeling the test molecule and/or the target molecule, where the label is covalently or non-covalently attached to the test molecule or target molecule. The label is sometimes a radioactive molecule such as  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{35}\text{S}$  or  $^3\text{H}$ , which can be detected by direct counting of radioemission or by scintillation counting. Also, enzymatic labels such as



horseradish peroxidase, alkaline phosphatase, or luciferase may be utilized where the enzymatic label can be detected by determining conversion of an appropriate substrate to product. In addition, presence or absence of an interaction can be determined without labeling. For example, a microphysiometer (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indication of an interaction between a test molecule and target molecule (McConnell, H. M. *et al.*, *Science* 257: 1906-1912 (1992)).

[0145] In cell-based systems, cells typically include a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleotide sequence referenced in Table A, an encoded polypeptide, or substantially identical nucleic acid or polypeptide thereof, and are often of mammalian origin, although the cell can be of any origin. Whole cells, cell homogenates, and cell fractions (*e.g.*, cell membrane fractions) can be subjected to analysis. Where interactions between a test molecule with a target polypeptide are monitored, soluble and/or membrane bound forms of the polypeptide may be utilized. Where membrane-bound forms of the polypeptide are used, it may be desirable to utilize a solubilizing agent. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecylpoly(ethylene glycol ether)<sub>n</sub>, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

[0146] An interaction between a test molecule and target molecule also can be detected by monitoring fluorescence energy transfer (FET) (*see, e.g.*, Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos *et al.* U.S. Patent No. 4,868,103). A fluorophore label on a first, "donor" molecule is selected such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, "acceptor" molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the "donor" polypeptide molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the "acceptor" molecule label may be differentiated from that of the "donor". Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, the spatial relationship between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the "acceptor" molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter).

[0147] In another embodiment, determining the presence or absence of an interaction between a test molecule and a target molecule can be effected by monitoring surface plasmon resonance (*see, e.g.*, Sjolander & Urbanicz, *Anal. Chem.* 63: 2338-2345 (1991) and Szabo *et al.*, *Curr. Opin. Struct. Biol.* 5:

699-705 (1995)). "Surface plasmon resonance" or "biomolecular interaction analysis (BIA)" can be utilized to detect biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

[0148] In another embodiment, the target molecule or test molecules are anchored to a solid phase, facilitating the detection of target molecule/test molecule complexes and separation of the complexes from free, uncomplexed molecules. The target molecule or test molecule is immobilized to the solid support. In an embodiment, the target molecule is anchored to a solid surface, and the test molecule, which is not anchored, can be labeled, either directly or indirectly, with detectable labels discussed herein.

[0149] It may be desirable to immobilize a target molecule, an anti-target molecule antibody, and/or test molecules to facilitate separation of target molecule/test molecule complexes from uncomplexed forms, as well as to accommodate automation of the assay. The attachment between a test molecule and/or target molecule and the solid support may be covalent or non-covalent (*see, e.g.*, U.S. Patent No. 6,022,688 for non-covalent attachments). The solid support may be one or more surfaces of the system, such as one or more surfaces in each well of a microtiter plate, a surface of a silicon wafer, a surface of a bead (*see, e.g.*, Lam, *Nature* 354: 82-84 (1991)) that is optionally linked to another solid support, or a channel in a microfluidic device, for example. Types of solid supports, linker molecules for covalent and non-covalent attachments to solid supports, and methods for immobilizing nucleic acids and other molecules to solid supports are well known (*see, e.g.*, U.S. Patent Nos. 6,261,776; 5,900,481; 6,133,436; and 6,022,688; and WIPO publication WO 01/18234).

[0150] In an embodiment, target molecule may be immobilized to surfaces via biotin and streptavidin. For example, biotinylated target polypeptide can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In another embodiment, a target polypeptide can be prepared as a fusion polypeptide. For example, glutathione-S-transferase/target polypeptide fusion can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivitized microtiter plates, which are then combined with a test molecule under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, or the matrix is immobilized in the case of beads, and complex formation is determined directly or indirectly as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of target molecule binding or activity is determined using standard techniques.

[0151] In an embodiment, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (e.g., by washing) under conditions such that a significant percentage of complexes formed will remain immobilized to the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of manners. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface, e.g., by adding a labeled antibody specific for the immobilized component, where the antibody, in turn, can be directly labeled or indirectly labeled with, e.g., a labeled anti-Ig antibody.

[0152] In another embodiment, an assay is performed utilizing antibodies that specifically bind target molecule or test molecule but do not interfere with binding of the target molecule to the test molecule. Such antibodies can be derivitized to a solid support, and unbound target molecule may be immobilized by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

[0153] Cell free assays also can be conducted in a liquid phase. In such an assay, reaction products are separated from unreacted components, by any of a number of standard techniques, including but not limited to: differential centrifugation (see, e.g., Rivas, G., and Minton, *Trends Biochem Sci Aug; 18(8): 284-7 (1993)*); chromatography (gel filtration chromatography, ion-exchange chromatography); electrophoresis (see, e.g., Ausubel et al., eds. *Current Protocols in Molecular Biology*, J. Wiley: New York (1999)); and immunoprecipitation (see, e.g., Ausubel et al., eds., *supra*). Media and chromatographic techniques are known to one skilled in the art (see, e.g., Heegaard, *J Mol. Recognit. Winter; 11(1-6): 141-8 (1998)*; Hage & Tweed, *J. Chromatogr. B Biomed. Sci. Appl. Oct 10; 699 (1-2): 499-525 (1997)*). Further, fluorescence energy transfer may also be conveniently utilized, as described herein, to detect binding without further purification of the complex from solution.

[0154] In another embodiment, modulators of target molecule expression are identified. For example, a cell or cell free mixture is contacted with a candidate compound and the expression of target mRNA or target polypeptide is evaluated relative to the level of expression of target mRNA or target polypeptide in the absence of the candidate compound. When expression of target mRNA or target polypeptide is greater in the presence of the candidate compound than in its absence, the candidate compound is identified as an agonist of target mRNA or target polypeptide expression. Alternatively, when expression of target mRNA or target polypeptide is less (e.g., less with statistical significance) in the presence of the candidate compound than in its absence, the candidate compound is identified as an

antagonist or inhibitor of target mRNA or target polypeptide expression. The level of target mRNA or target polypeptide expression can be determined by methods described herein.

[0155] In another embodiment, binding partners that interact with a target molecule are detected. The target molecules can interact with one or more cellular or extracellular macromolecules, such as polypeptides *in vivo*, and these interacting molecules are referred to herein as “binding partners.” Binding partners can agonize or antagonize target molecule biological activity. Also, test molecules that agonize or antagonize interactions between target molecules and binding partners can be useful as therapeutic molecules as they can up-regulate or down-regulated target molecule activity *in vivo* and thereby treat osteoarthritis.

[0156] Binding partners of target molecules can be identified by methods known in the art. For example, binding partners may be identified by lysing cells and analyzing cell lysates by electrophoretic techniques. Alternatively, a two-hybrid assay or three-hybrid assay can be utilized (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.*, *Cell* 72:223-232 (1993); Madura *et al.*, *J. Biol. Chem.* 268: 12046-12054 (1993); Bartel *et al.*, *Biotechniques* 14: 920-924 (1993); Iwabuchi *et al.*, *Oncogene* 8: 1693-1696 (1993); and Brent WO94/10300). A two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. The assay often utilizes two different DNA constructs. In one construct, a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A (sometimes referred to as the “bait”) is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In another construct, a DNA sequence from a library of DNA sequences that encodes a potential binding partner (sometimes referred to as the “prey”) is fused to a gene that encodes an activation domain of the known transcription factor. Sometimes, a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A can be fused to the activation domain. If the “bait” and the “prey” molecules interact *in vivo*, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to identify the potential binding partner.

[0157] In an embodiment for identifying test molecules that antagonize or agonize complex formation between target molecules and binding partners, a reaction mixture containing the target molecule and the binding partner is prepared, under conditions and for a time sufficient to allow complex formation. The reaction mixture often is provided in the presence or absence of the test molecule. The test molecule can be included initially in the reaction mixture, or can be added at a time subsequent to the addition of the target molecule and its binding partner. Control reaction mixtures are incubated without the test molecule or with a placebo. Formation of any complexes between the target molecule and the

binding partner then is detected. Decreased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule antagonizes target molecule/binding partner complex formation. Alternatively, increased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule agonizes target molecule/binding partner complex formation. In another embodiment, complex formation of target molecule/binding partner can be compared to complex formation of mutant target molecule/binding partner (*e.g.*, amino acid modifications in a target polypeptide). Such a comparison can be important in those cases where it is desirable to identify test molecules that modulate interactions of mutant but not non-mutated target gene products.

[0158] The assays can be conducted in a heterogeneous or homogeneous format. In heterogeneous assays, target molecule and/or the binding partner are immobilized to a solid phase, and complexes are detected on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the molecules being tested. For example, test compounds that agonize target molecule/binding partner interactions can be identified by conducting the reaction in the presence of the test molecule in a competition format. Alternatively, test molecules that agonize preformed complexes, *e.g.*, molecules with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed.

[0159] In a heterogeneous assay embodiment, the target molecule or the binding partner is anchored onto a solid surface (*e.g.*, a microtiter plate), while the non-anchored species is labeled, either directly or indirectly. The anchored molecule can be immobilized by non-covalent or covalent attachments. Alternatively, an immobilized antibody specific for the molecule to be anchored can be used to anchor the molecule to the solid surface. The partner of the immobilized species is exposed to the coated surface with or without the test molecule. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) such that a significant portion of any complexes formed will remain immobilized on the solid surface. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface is indicative of complex. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored to the surface; *e.g.*, by using a labeled antibody specific for the initially non-immobilized species. Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

[0160] In another embodiment, the reaction can be conducted in a liquid phase in the presence or absence of test molecule, where the reaction products are separated from unreacted components, and the complexes are detected (*e.g.*, using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to

detect anchored complexes). Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex or that disrupt preformed complexes can be identified.

[0161] In an alternate embodiment, a homogeneous assay can be utilized. For example, a preformed complex of the target gene product and the interactive cellular or extracellular binding partner product is prepared. One or both of the target molecule or binding partner is labeled, and the signal generated by the label(s) is quenched upon complex formation (e.g., U.S. Patent No. 4,109,496 that utilizes this approach for immunoassays). Addition of a test molecule that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt target molecule/binding partner complexes can be identified.

[0162] Candidate therapeutics for treating osteoarthritis are identified from a group of test molecules that interact with a target molecule. Test molecules are normally ranked according to the degree with which they modulate (e.g., agonize or antagonize) a function associated with the target molecule (e.g., DNA replication and/or processing, RNA transcription and/or processing, polypeptide production and/or processing, and/or biological function/activity), and then top ranking modulators are selected. Also, pharmacogenomic information described herein can determine the rank of a modulator. The top 10% of ranked test molecules often are selected for further testing as candidate therapeutics, and sometimes the top 15%, 20%, or 25% of ranked test molecules are selected for further testing as candidate therapeutics. Candidate therapeutics typically are formulated for administration to a subject.

#### Therapeutic Formulations

[0163] Formulations and pharmaceutical compositions typically include in combination with a pharmaceutically acceptable carrier one or more target molecule modulators. The modulator often is a test molecule identified as having an interaction with a target molecule by a screening method described above. The modulator may be a compound, an antisense nucleic acid, a ribozyme, an antibody, or a binding partner. Also, formulations may comprise a target polypeptide or fragment thereof in combination with a pharmaceutically acceptable carrier.

[0164] As used herein, the term "pharmaceutically acceptable carrier" includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions. Pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

[0165] A pharmaceutical composition typically is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application

can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerin, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0166] Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, *e.g.*, gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[0167] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[0168] Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as

required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0169] For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

[0170] Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art. Molecules can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

[0171] In one embodiment, active molecules are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. Materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

[0172] It is advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier.

[0173] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the



population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio  $LD_{50}/ED_{50}$ . Molecules which exhibit high therapeutic indices are preferred. While molecules that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0174] The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such molecules lies preferably within a range of circulating concentrations that include the  $ED_{50}$  with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any molecules used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the  $IC_{50}$  (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

[0175] As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, sometimes about 0.01 to 25 mg/kg body weight, often about 0.1 to 20 mg/kg body weight, and more often about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The protein or polypeptide can be administered one time per week for between about 1 to 10 weeks, sometimes between 2 to 8 weeks, often between about 3 to 7 weeks, and more often for about 4, 5, or 6 weeks. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

[0176] With regard to polypeptide formulations, featured herein is a method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject with a first polypeptide, where the subject comprises a second polypeptide having one or more polymorphic variations associated with cancer, and where the first polypeptide comprises fewer polymorphic variations associated with cancer than the second polypeptide. The first and second polypeptides are encoded by a nucleic acid which comprises a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A; a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence referenced in SEQ ID NO: 1-3 or referenced in Table A; a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded

by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A and a nucleotide sequence 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A. The subject often is a human.

[0177] For antibodies, a dosage of 0.1 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg) is often utilized. If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is often appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (e.g., into the brain). A method for lipidation of antibodies is described by Cruikshank *et al.*, *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193 (1997).

[0178] Antibody conjugates can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

[0179] For compounds, exemplary doses include milligram or microgram amounts of the compound per kilogram of subject or sample weight, for example, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. When one or more of these small molecules is to be administered to an animal (e.g., a human) in order to modulate expression or activity of a polypeptide or nucleic acid described herein, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

[0180] With regard to nucleic acid formulations, gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.*, U.S. Patent 5,328,470) or by stereotactic injection (*see e.g.*, Chen *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). Pharmaceutical preparations of gene therapy vectors can include a gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells (*e.g.*, retroviral vectors) the pharmaceutical preparation can include one or more cells which produce the gene delivery system. Examples of gene delivery vectors are described herein.

#### Therapeutic Methods

[0181] A therapeutic formulation described above can be administered to a subject in need of a therapeutic for inducing a desired biological response. Therapeutic formulations can be administered by any of the paths described herein. With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from pharmacogenomic analyses described herein.

[0182] As used herein, the term "treatment" is defined as the application or administration of a therapeutic formulation to a subject, or application or administration of a therapeutic agent to an isolated tissue or cell line from a subject with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect osteoarthritis, symptoms of osteoarthritis or a predisposition towards osteoarthritis. A therapeutic formulation includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. Administration of a therapeutic formulation can occur prior to the manifestation of symptoms characteristic of osteoarthritis, such that osteoarthritis is prevented or delayed in its progression. The appropriate therapeutic composition can be determined based on screening assays described herein.

[0183] As discussed, successful treatment of osteoarthritis can be brought about by techniques that serve to agonize target molecule expression or function, or alternatively, antagonize target molecule expression or function. These techniques include administration of modulators that include, but are not limited to, small organic or inorganic molecules; antibodies (including, for example, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and Fab, F(ab')<sub>2</sub> and Fab expression library fragments, scFV molecules, and epitope-binding fragments thereof); and peptides, phosphopeptides, or polypeptides.

[0184] Further, antisense and ribozyme molecules that inhibit expression of the target gene can also be used to reduce the level of target gene expression, thus effectively reducing the level of target gene activity. Still further, triple helix molecules can be utilized in reducing the level of target gene activity. Antisense, ribozyme and triple helix molecules are discussed above. It is possible that the use of

antisense, ribozyme, and/or triple helix molecules to reduce or inhibit mutant gene expression can also reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles, such that the concentration of normal target gene product present can be lower than is necessary for a normal phenotype. In such cases, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity can be introduced into cells via gene therapy method. Alternatively, in instances in that the target gene encodes an extracellular polypeptide, it can be preferable to co-administer normal target gene polypeptide into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

[0185] Another method by which nucleic acid molecules may be utilized in treating or preventing osteoarthritis is use of aptamer molecules specific for target molecules. Aptamers are nucleic acid molecules having a tertiary structure which permits them to specifically bind to ligands (*see, e.g., Osborne, et al., Curr. Opin. Chem. Biol.* 1(1): 5-9 (1997); and Patel, D. J., *Curr. Opin. Chem. Biol. Jun;1(1): 32-46 (1997)*).

[0186] Yet another method of utilizing nucleic acid molecules for osteoarthritis treatment is gene therapy, which can also be referred to as allele therapy. Provided herein is a gene therapy method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject or from the subject with a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-3 or other nucleotide sequence referenced in Table A). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-3 or other nucleotide sequence referenced in Table A). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human. Allele therapy methods often are utilized in conjunction with a method of first determining whether a subject has genomic DNA that includes polymorphic variants associated with osteoarthritis.

[0187] In another allele therapy embodiment, provided herein is a method which comprises contacting one or more cells in the subject or from the subject with a polypeptide encoded by a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to the nucleotide sequence of SEQ ID NO: 1-3 or other nucleotide sequence referenced in Table A). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-3 or other nucleotide sequence

referenced in Table A). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human.

[0188] For antibody-based therapies, antibodies can be generated that are both specific for target molecules and that reduce target molecule activity. Such antibodies may be administered in instances where antagonizing a target molecule function is appropriate for the treatment of osteoarthritis.

[0189] In circumstances where stimulating antibody production in an animal or a human subject by injection with a target molecule is harmful to the subject, it is possible to generate an immune response against the target molecule by use of anti-idiotypic antibodies (*see, e.g., Herlyn, Ann. Med.;31(1): 66-78 (1999); and Bhattacharya-Chatterjee & Foon, Cancer Treat. Res.; 94: 51-68 (1998)*). Introducing an anti-idiotypic antibody to a mammal or human subject often stimulates production of anti-anti-idiotypic antibodies, which typically are specific to the target molecule. Vaccines directed to osteoarthritis also may be generated in this fashion.

[0190] In instances where the target molecule is intracellular and whole antibodies are used, internalizing antibodies may be preferred. Lipofectin or liposomes can be used to deliver the antibody or a fragment of the Fab region that binds to the target antigen into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target antigen is preferred. For example, peptides having an amino acid sequence corresponding to the Fv region of the antibody can be used. Alternatively, single chain neutralizing antibodies that bind to intracellular target antigens can also be administered. Such single chain antibodies can be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population (*see, e.g., Marasco et al., Proc. Natl. Acad. Sci. USA 90: 7889-7893 (1993)*).

[0191] Modulators can be administered to a patient at therapeutically effective doses to treat osteoarthritis. A therapeutically effective dose refers to an amount of the modulator sufficient to result in amelioration of symptoms of osteoarthritis. Toxicity and therapeutic efficacy of modulators can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g., for determining the LD<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population)*. The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD<sub>50</sub>/ED<sub>50</sub>. Modulators that exhibit large therapeutic indices are preferred. While modulators that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such molecules to the site of affected tissue in order to minimize potential damage to uninfected cells, thereby reducing side effects.

[0192] Data obtained from cell culture assays and animal studies can be used in formulating a range of dosages for use in humans. The dosage of such compounds lies preferably within a range of

circulating concentrations that include the  $ED_{50}$  with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the  $IC_{50}$  (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma can be measured, for example, by high performance liquid chromatography.

[0193] Another example of effective dose determination for an individual is the ability to directly assay levels of “free” and “bound” compound in the serum of the test subject. Such assays may utilize antibody mimics and/or “biosensors” that have been created through molecular imprinting techniques. Molecules that modulate target molecule activity are used as a template, or “imprinting molecule”, to spatially organize polymerizable monomers prior to their polymerization with catalytic reagents. The subsequent removal of the imprinted molecule leaves a polymer matrix which contains a repeated “negative image” of the compound and is able to selectively rebind the molecule under biological assay conditions. A detailed review of this technique can be seen in Ansell *et al.*, *Current Opinion in Biotechnology* 7: 89-94 (1996) and in Shea, *Trends in Polymer Science* 2: 166-173 (1994). Such “imprinted” affinity matrixes are amenable to ligand-binding assays, whereby the immobilized monoclonal antibody component is replaced by an appropriately imprinted matrix. An example of the use of such matrixes in this way can be seen in Vlatakis, *et al.*, *Nature* 361: 645-647 (1993). Through the use of isotope-labeling, the “free” concentration of compound which modulates target molecule expression or activity readily can be monitored and used in calculations of  $IC_{50}$ . Such “imprinted” affinity matrixes can also be designed to include fluorescent groups whose photon-emitting properties measurably change upon local and selective binding of target compound. These changes readily can be assayed in real time using appropriate fiberoptic devices, in turn allowing the dose in a test subject to be quickly optimized based on its individual  $IC_{50}$ . An example of such a “biosensor” is discussed in Kriz *et al.*, *Analytical Chemistry* 67: 2142-2144 (1995).

[0194] The examples set forth below are intended to illustrate but not limit the invention.

### Examples

[0195] In the following studies a group of subjects was selected according to specific parameters pertaining to osteoarthritis. Nucleic acid samples obtained from individuals in the study group were subjected to genetic analyses that identified associations between osteoarthritis and certain polymorphic variants in human genomic DNA. The polymorphisms were genotyped again in two replication cohorts consisting of individuals selected for OA. In addition, SNPs proximal to the incident polymorphism in

the *KIAA0296* region, the *SNW1* region and in chromosome 4 were identified and allelotyped in OA case and control pools. Methods are described for producing target polypeptides encoded by the nucleic acids of Table A *in vitro* or *in vivo*, which can be utilized in methods that screen test molecules for those that interact with target polypeptides. Test molecules identified as being interactors with target polypeptides can be screened further as osteoarthritis therapeutics.

### Example 1

#### Samples and Pooling Strategies

##### Sample Selection

[0196] Blood samples were collected from individuals diagnosed with knee osteoarthritis, which were referred to as case samples. Also, blood samples were collected from individuals not diagnosed with knee osteoarthritis as gender and age-matched controls. A database was created that listed all phenotypic trait information gathered from individuals for each case and control sample. Genomic DNA was extracted from each of the blood samples for genetic analyses.

##### DNA Extraction from Blood Samples

[0197] Six to ten milliliters of whole blood was transferred to a 50 ml tube containing 27 ml of red cell lysis solution (RCL). The tube was inverted until the contents were mixed. Each tube was incubated for 10 minutes at room temperature and inverted once during the incubation. The tubes were then centrifuged for 20 minutes at 3000 x g and the supernatant was carefully poured off. 100-200 µl of residual liquid was left in the tube and was pipetted repeatedly to resuspend the pellet in the residual supernatant. White cell lysis solution (WCL) was added to the tube and pipetted repeatedly until completely mixed. While no incubation was normally required, the solution was incubated at 37°C or room temperature if cell clumps were visible after mixing until the solution was homogeneous. 2 ml of protein precipitation was added to the cell lysate. The mixtures were vortexed vigorously at high speed for 20 sec to mix the protein precipitation solution uniformly with the cell lysate, and then centrifuged for 10 minutes at 3000 x g. The supernatant containing the DNA was then poured into a clean 15 ml tube, which contained 7 ml of 100% isopropanol. The samples were mixed by inverting the tubes gently until white threads of DNA were visible. Samples were centrifuged for 3 minutes at 2000 x g and the DNA was visible as a small white pellet. The supernatant was decanted and 5 ml of 70% ethanol was added to each tube. Each tube was inverted several times to wash the DNA pellet, and then centrifuged for 1 minute at 2000 x g. The ethanol was decanted and each tube was drained on clean absorbent paper. The DNA was dried in the tube by inversion for 10 minutes, and then 1000 µl of 1X TE was added. The size of each sample was estimated, and less TE buffer was added during the following DNA hydration step if

the sample was smaller. The DNA was allowed to rehydrate overnight at room temperature, and DNA samples were stored at 2-8°C.

[0198] DNA was quantified by placing samples on a hematology mixer for at least 1 hour. DNA was serially diluted (typically 1:80, 1:160, 1:320, and 1:640 dilutions) so that it would be within the measurable range of standards. 125 µl of diluted DNA was transferred to a clear U-bottom microtitre plate, and 125 µl of 1X TE buffer was transferred into each well using a multichannel pipette. The DNA and 1X TE were mixed by repeated pipetting at least 15 times, and then the plates were sealed. 50 µl of diluted DNA was added to wells A5-H12 of a black flat bottom microtitre plate. Standards were inverted six times to mix them, and then 50 µl of 1X TE buffer was pipetted into well A1, 1000 ng/ml of standard was pipetted into well A2, 500 ng/ml of standard was pipetted into well A3, and 250 ng/ml of standard was pipetted into well A4. PicoGreen (Molecular Probes, Eugene, Oregon) was thawed and freshly diluted 1:200 according to the number of plates that were being measured. PicoGreen was vortexed and then 50µl was pipetted into all wells of the black plate with the diluted DNA. DNA and PicoGreen were mixed by pipetting repeatedly at least 10 times with the multichannel pipette. The plate was placed into a Fluoroskan Ascent Machine (microplate fluorometer produced by Labsystems) and the samples were allowed to incubate for 3 minutes before the machine was run using filter pairs 485 nm excitation and 538 nm emission wavelengths. Samples having measured DNA concentrations of greater than 450 ng/µl were re-measured for conformation. Samples having measured DNA concentrations of 20 ng/µl or less were re-measured for confirmation.

#### Pooling Strategies – Discovery Cohort

[0199] Samples were derived from the Nottingham knee OA family study (UK) where index cases were identified through a knee replacement registry. Siblings were approached and assessed with knee x-rays and assigned status as affected or unaffected. In all 1,157 individuals were available. In order to create same-sex pools of appropriate sizes, 335 unrelated female individuals with OA from the Nottingham OA sample were selected for the case pool. The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The female case samples and female control samples are described further in Table 1 below.

[0200] A select set of samples from each group were utilized to generate pools, and one pool was created for each group. Each individual sample in a pool was represented by an equal amount of genomic DNA. For example, where 25 ng of genomic DNA was utilized in each PCR reaction and there



were 200 individuals in each pool, each individual would provide 125 pg of genomic DNA. Inclusion or exclusion of samples for a pool was based upon the following criteria: the sample was derived from an individual characterized as Caucasian; the sample was derived from an individual of British paternal and maternal descent; case samples were derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic. Control samples were derived from individuals free of OA, family history of OA, and rheumatoid arthritis. Also, sufficient genomic DNA was extracted from each blood sample for all allelotyping and genotyping reactions performed during the study. Phenotype information from each individual was collected and included age of the individual, gender, family history of OA, general medical information (e.g., height, weight, thyroid disease, diabetes, psoriasis, hysterectomy), joint history (previous and current symptoms, joint-related operations, age at onset of symptoms, date of primary diagnosis, age of individual as of primary diagnosis and order of involvement), and knee-related findings (crepitus, restricted passive movement, bony swelling/deformity). Additional knee information included knee history, current symptoms, any major knee injury, meniscectomy, knee replacement surgery, age of surgery, and treatment history (including hormone replace therapy (HRT)). Samples that met these criteria were added to appropriate pools based on disease status.

[0201] The selection process yielded the pools set forth in Table 1, which were used in the studies that follow:

**TABLE 1**

	<b>Female case</b>	<b>Female control</b>
<b>Pool size (Number)</b>	335	335
<b>Pool Criteria (ex: case/control)</b>	control	case
<b>Mean Age (ex: years)</b>	57.21	69.95

### Example 2

#### Association of Polymorphic Variants with Osteoarthritis

[0202] A whole-genome screen was performed to identify particular SNPs associated with occurrence of osteoarthritis. As described in Example 1, two sets of samples were utilized, which included samples from female individuals having knee osteoarthritis (osteoarthritis cases), and samples from female individuals not having knee osteoarthritis (female controls). The initial screen of each pool was performed in an allelotyping study, in which certain samples in each group were pooled. By pooling DNA from each group, an allele frequency for each SNP in each group was calculated. These allele

frequencies were then compared to one another. Particular SNPs were considered as being associated with osteoarthritis when allele frequency differences calculated between case and control pools were statistically significant. SNP disease association results obtained from the allelotyping study were then validated by genotyping each associated SNP across all samples from each pool. The results of the genotyping then were analyzed, allele frequencies for each group were calculated from the individual genotyping results, and a p-value was calculated to determine whether the case and control groups had statistically significant differences in allele frequencies for a particular SNP. When the genotyping results agreed with the original allelotyping results, the SNP disease association was considered validated at the genetic level.

#### SNP Panel Used for Genetic Analyses

[0203] A whole-genome SNP screen began with an initial screen of approximately 25,000 SNPs over each set of disease and control samples using a pooling approach. The pools studied in the screen are described in Example 1. The SNPs analyzed in this study were part of a set of 25,488 SNPs confirmed as being statistically polymorphic as each is characterized as having a minor allele frequency of greater than 10%. The SNPs in the set reside in genes or in close proximity to genes, and many reside in gene exons. Specifically, SNPs in the set are located in exons, introns, and within 5,000 base-pairs upstream of a transcription start site of a gene. In addition, SNPs were selected according to the following criteria: they are located in ESTs; they are located in Locuslink or Ensembl genes; and they are located in Genomatix promoter predictions. SNPs in the set were also selected on the basis of even spacing across the genome, as depicted in Table 2.

[0204] A case-control study design using a whole genome association strategy involving approximately 28,000 single nucleotide polymorphisms (SNPs) was employed. Approximately 25,000 SNPs were evenly spaced in gene-based regions of the human genome with a median inter-marker distance of about 40,000 base pairs. Additionally, approximately 3,000 SNPs causing amino acid substitutions in genes described in the literature as candidates for various diseases were used. The case-control study samples were of female Caucasian origin (British paternal and maternal descent) 670 individuals were equally distributed in two groups: female controls and female cases. The whole genome association approach was first conducted on 2 DNA pools representing the 2 groups. Significant markers were confirmed by individual genotyping.

TABLE 2

General Statistics		Spacing Statistics	
Total # of SNPs	25,488	Median	37,058 bp
# of Exonic SNPs	>4,335 (17%)	Minimum*	1,000 bp
# SNPs with refSNP ID	20,776 (81%)	Maximum*	3,000,000 bp
Gene Coverage	>10,000	Mean	122,412 bp
Chromosome Coverage	All	Std Deviation	373,325 bp
		*Excludes outliers	

#### Allelotyping and Genotyping Results

[0205] The genetic studies summarized above and described in more detail below identified allelic variants associated with osteoarthritis, which are summarized in Table A.

#### Assay for Verifying, Allelotyping, and Genotyping SNPs

[0206] A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hMET™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0207] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which were used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension primers used for analyzing polymorphisms. The initial PCR amplification reaction was performed in a 5 µl total volume containing 1X PCR buffer with 1.5 mM MgCl<sub>2</sub> (Qiagen), 200 µM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

TABLE 3: PCR Primers

SNP Reference	Forward PCR primer	Reverse PCR primer
rs552	ACGTTGGATGGACTGAGGTAGATGATGC	ACGTTGGATGGCTTTCCTTCCCTTGGTTTC
rs12904	ACGTTGGATGGAACCACTCCACCACAG	ACGTTGGATGGGTGGGGATGGCACTGTC
rs2282146	ACGTTGGATGTCCACGAGGACCTGGAG	ACGTTGGATGTTCTGTTGGGTGGCCGGG

SNP Reference	Forward PCR primer	Reverse PCR primer
rs734784	ACGTTGGATGTCGGGATGTCTCCAGAGATG	ACGTTGGATGGCAACCAACCAAGATTTGAG
rs1042164	ACGTTGGATGTTTCTTCCAGACGGGCTTTC	ACGTTGGATGCAAAAGTCAGCCGCCAAACGAC
rs749670	ACGTTGGATGTCCTCATCTGTGTGCCATTG	ACGTTGGATGATBAGGGGTGAAAGCCAGGAG
rs955592	ACGTTGGATGTTCCCATCTTCTCTGGGCTC	ACGTTGGATGTCTCAGAGGGGTCTCCTTTTC
rs1143016	ACGTTGGATGTTGTCCAGCAGGTAGGGCAG	ACGTTGGATGACCCATCGCGGATACATGTG
rs755248	ACGTTGGATGGGCTCTCTGCTGAGGAAGTGG	ACGTTGGATGACATCCTACTCAGCGGGCCAG
rs1055055	ACGTTGGATGTTGTGCTTCTGAGGAATCC	ACGTTGGATGGTGCAGAGAGCGTCTATAC
rs835409	ACGTTGGATGCTCTGTTGGCTTTTGCAGAC	ACGTTGGATGACTCTCATGGTGGTGAAC
rs927663	ACGTTGGATGTTTGACTGGTGGCCCAAC	ACGTTGGATGAAGAACTCTCAGTGCCAGCC
rs8162	ACGTTGGATGCTTCATCCAGAACCTCCAGG	ACGTTGGATGTGCATATGGCTGTGCAGAGC
rs831038	ACGTTGGATGTGAAAGAGCTGCCCTCTTTTC	ACGTTGGATGAAATGACACTCAGCGGTAAAG
rs33079	ACGTTGGATGTTATTTCATTGGCCAGGCC	ACGTTGGATGGTGTCTCAGTGTCTATGAC
rs1710880	ACGTTGGATGCGAAGGCCAGAAATAAAGTG	ACGTTGGATGAACCTCTGTGTTTAAAGAAC
rs1078153	ACGTTGGATGTCTGCGTGTAACTGAGAGG	ACGTTGGATGAACATACACACAGTGCGAGC
rs799570	ACGTTGGATGATGCATATGGGCAAGTTGCC	ACGTTGGATGCCAGGAAAGCATCCTCAGAC
rs1282730	ACGTTGGATGGGCTCTTGTACTGTGCTG	ACGTTGGATGACGAAAAGAGGTTGTGTACAG
rs1518875	ACGTTGGATGAGAATGCGTTCAATGCGCTGC	ACGTTGGATGAGCGAAAAGCTCTGCCATTG
rs1568694	ACGTTGGATGTTTCACTCAGTTATGGACGG	ACGTTGGATGAACCTCTGTGTTTAAAGAAC
rs905042	ACGTTGGATGTACAATGGTAAGGGCCAAAG	ACGTTGGATGGGTCCATAATGGTCAATTGTG
rs1957723	ACGTTGGATGTACTCACTTGTGTACTGCTC	ACGTTGGATGGCTGCAAGCGTCACATTAATC
rs794018	ACGTTGGATGGGATGATGATGAAATGACTG	ACGTTGGATGGCTCTGATTTAGATGAGTCTC
rs707723	ACGTTGGATGTGTGGCTGAAGTTTGTCTGCT	ACGTTGGATGCACACACAAACCTTTGAAGAG
rs893861	ACGTTGGATGGAGGCATGTACACAAAAGTG	ACGTTGGATGCTCAGCACTGTAATAGTTG
rs1914903	ACGTTGGATGTGCGTCAAGTTGAAGTCTCT	ACGTTGGATGAGGGTAGTGAGTTCACATGC
rs2062232	ACGTTGGATGTCTGCTCAGATAAAGTGCTG	ACGTTGGATGGCGGTAGTGTTCCTTAAACC
rs26809	ACGTTGGATGCAATGGGAGATCAGAAACATC	ACGTTGGATGAACTTCAATGTTCTGTATGAG
rs1370987	ACGTTGGATGATACTTTGGATGTCTGGTGG	ACGTTGGATGGGTCTTGGTCAACACTATC
rs1012414	ACGTTGGATGACTTGGAAAGTCAGTCTGGG	ACGTTGGATGGAAACCGAGAAATGCGTATG
rs435903	ACGTTGGATGGGCATAAGTTAGAGACAACC	ACGTTGGATGGGCTATGTATGCTGCTGTG
rs1248	ACGTTGGATGGAGATTGTGCATTTGGCCAAG	ACGTTGGATGCAGACCAACCTTTAACCAGG
rs703508	ACGTTGGATGAGCTCTGTGGCTCTTTTGG	ACGTTGGATGTACTCACAGCTTCCCGGCG
rs226465	ACGTTGGATGAATTTTGACCCTGCCAACCC	ACGTTGGATGTATGTGAAAGAGGCGTGAAG
rs241448	ACGTTGGATGCAAGCTGCAGAAAGCTTGCC	ACGTTGGATGTGAGAAGAGGGCCAGATATC
rs763155	ACGTTGGATGGGGGAAACCAAAATAGTGTG	ACGTTGGATGTACAGGAGAGTAAATGGCTC
rs1040461	ACGTTGGATGACATCTGGTGGAAAGTCACTC	ACGTTGGATGGGTCTTGTGTTGTGGGCTC
rs462832	ACGTTGGATGCACTTCTCTGTAATATTG	ACGTTGGATGTGAGACACAAATAATTTGCC
rs804194	ACGTTGGATGTAATCCGGTGGCAGATCAAG	ACGTTGGATGGAATTCATGTGCTGACGGG
rs1022646	ACGTTGGATGACTGTCACTAATCATCTGTG	ACGTTGGATGGACTATGTTGGATGTCAGAG
rs1569112	ACGTTGGATGTCATGATGCTGCGTGTGGAGA	ACGTTGGATGACCATCTCACACCCATGCA
rs805623	ACGTTGGATGATCAACCACTCATACACTGCT	ACGTTGGATGCACAGAAACAGCTGGAATGC
rs1019850	ACGTTGGATGTTTACTCTCAGGAAGCACCC	ACGTTGGATGACGAGGAGGAATGTTCCAG
rs1589931	ACGTTGGATGTCAAAACCTTCTCCTGTAGACG	ACGTTGGATGTGAACATAGTAGGCCGCTCAG
AA	ACGTTGGATGTAGAGTGCTCGTATTTTGG	ACGTTGGATGCTGGGAACAGCTTTTGTATCC
rs279941	ACGTTGGATGTCATAGGGAACACCGAGAATG	ACGTTGGATGGGTGTTCATCTATGGGCTAG
rs1062230	ACGTTGGATGAAACTCCTTTCCCTCTCAAAG	ACGTTGGATGGGCGCATCAGTCTATAGTTT
rs1859911	ACGTTGGATGCTGTTTTCGAGCATCTAC	ACGTTGGATGCTCTTGCATATGAGATAGG
rs1477261	ACGTTGGATGCAGGGTTATGTGGTATTATC	ACGTTGGATGGGGAAGATGAAAGATAAGAG
rs1191119	ACGTTGGATGACTCTCAGGGTGATTATCTG	ACGTTGGATGTGTAAGATTCTGGCACTGTG
rs657780	ACGTTGGATGTTTAAAGAACGCCCAAGGAG	ACGTTGGATGTCATGTTTCAAGCACTTTGG
rs1393890	ACGTTGGATGGTCTGATTATCTTCTGCCG	ACGTTGGATGGGTACCTTTATCCTTGCTTC
rs1478714	ACGTTGGATGAATAATTTGCTGACACCCGCC	ACGTTGGATGGGAGTCCAGAGGTTTAAACAG

SNP Reference	Forward PCR primer	Reverse PCR primer
rs868213	ACGTTGGATGTGTCAGAACTGGGCACATTC	ACGTTGGATGAGGGATAGGGATCAGGAATG
rs690115	ACGTTGGATGAATAGCCAAGGCCGTGTGGG	ACGTTGGATGCACCTGGGAGATAGCAGGG
rs1485501	ACGTTGGATGTCAGGAATTGTTACCTGGAC	ACGTTGGATGCCCTCATCTAGACACTTTTG
rs899173	ACGTTGGATGAGTGCCACATCACTCTTGTG	ACGTTGGATGTTCTGCTCCACTACAGTCTC
rs10477	ACGTTGGATGGGGGTACGTGGAAGTTACC	ACGTTGGATGATGGCAATCAAGAGAGTCTAA
rs926393	ACGTTGGATGAGATCAGCCAGGAATGTG	ACGTTGGATGTGTTGGAGAAGGTTTCCACC
rs465271	ACGTTGGATGAATCAGAGCTCATGGCTCAC	ACGTTGGATGATGGTATGTGCACCTATGG
rs13847	ACGTTGGATGCCGCCGTAGTGATAAGCAC	ACGTTGGATGCAGGACAGGGCAGAGTGAG
rs738658	ACGTTGGATGGATGATGTGTGCATCAGG	ACGTTGGATGCTTCCAAGAGATGGCGTTC

[0208] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 µl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 µl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0209] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 4, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

TABLE 4: Extension Primers

SNP Reference	Extend Probe	Termination Mix
rs552	TGATGCTGTTGTCAGATACC	ACT
rs12904	AGCCTCAAACGGGTCA	ACT
rs2282146	GGACCTGGAGCCCCCACC	ACG
rs734784	GCCTCCGACACCCCATCAA	ACG
rs1042164	CTTGCTCGGGACCAAGTCCA	ACT
rs749670	GTTGGTGGGCATCCCTTTC	ACG
rs955592	TGGGCTCTGACCACCTCT	ACT
rs1143016	ATGCAGCGTCAACAGCAC	ACT
rs755248	TGAGGAAGTGCCAGGTGTG	ACG
rs1055055	CCGAGTTCAAGGCTCACTTTC	ACT
rs835409	TTCAGACCAAGCCAAATTAAGAA	ACT
rs927663	GGTTGCCCAAACTCCCTT	ACT
rs8162	AACACAGAGCAAAGCACC	ACT
rs831038	CGTTATAGTAAGGAAAGGCAG	ACG

SNP Reference	Extend Probe	Termination Mix
rs33079	CCCATCACCTGGAGCTTTG	ACG
rs1710880	CTGTATTATGTTCCCTTGG	CGT
rs1078153	GCCGGCACCGTCAGAAAC	CGT
rs799570	GCAGTTCCTAGAAGACAGCT	ACT
rs1282730	TGCTGGCCCAACTTTTGCT	ACG
rs1518875	CTGCAATGTTTCCAAACCCC	ACT
rs1568694	AGTTATGGACGGAAGAGGG	ACG
rs905042	GGTAAGGGCCAAGTGAGTG	CGT
rs1957723	AGCATGGCATAGGCACTGG	ACG
rs794018	AAATGACTGAAATGTGTACTATA	ACG
rs707723	CCTGAGGTATATTCAATA	ACG
rs893861	CATGTACACAAAACGTGTAAGTAA	ACG
rs1914903	TCCCCATAGATGGACCTGC	ACG
rs2062232	GCTGAAGACAAGGATTAGGTT	ACG
rs26609	GAGATCAGAAACATCACCTTG	CGT
rs1370987	TTTGGGAGTTACTGCCTTAGAA	ACT
rs1012414	ACTAGGAACCCAGAAATATGAGCATC	ACG
rs435903	AAGCTACAATGGAATAATGGC	ACG
rs1248	GTGCATTTTGGCAAGAATATATG	CGT
rs703508	GGGGTCCAGGCAGAAAGAAAC	ACG
rs226465	CCTCTTCCCCTCCTCCCT	ACT
rs241448	GCAGAAAGCTTGCCAGCTC	ACG
rs783155	CGACCGTCCAAGTGAGTGA	ACT
rs1040481	AAGTCACTCCGGTCAGAAATCA	ACT
rs462832	ATAAGAATCTTTTATGATCCCAACA	CGT
rs804194	GATCAAGGCTGATCTCGCC	ACT
rs1022846	CCTAATCATCCTGCCACCC	ACT
rs1569112	ACCAGGCCCATGGGCTG	ACG
rs805623	CTGTGTTCAAATAAGGCAACC	ACT
rs1019850	AGGAAGCCACCCAGCTAATAC	CGT
rs1599931	CTGAGGCCGGGAGGGATT	ACT
AA	TAGTTTTAAATTCTGCACA	ACT
rs279941	AACACCGAGAATGAAACATC	ACT
rs1062230	ATGCTGGTCTGTCCAA	ACG
rs1859911	TCCGAGCATCTACATGCTCA	ACT
rs1477261	AGGAGGAGCCCAATATGAAA	CGT
rs1191119	GTCTTTTGTAACTGGGGAACCC	ACG
rs657780	CGCCAAGGAGTTTCCACA	ACT
rs1393890	CTGCCGTACCTGGCAAGC	ACT
rs1478714	CCCCGAGGGGACAGTCCA	ACG
rs868213	GGGCACATTCTTGAGGAGGT	ACG
rs690115	AGCCGAGGGAGCTGACCCTG	ACG
rs1465501	TCCAGGAGCCCTCAGAAATG	ACT
rs899173	CCTCTGGCAAAGTGTGGAGC	ACG
rs10477	AGTACGATATCAAGATC	ACG
rs926393	CAGGAAATGTGCTTTGAGTTCC	ACG
rs465271	GGCTCAAGGGATCCTCCCA	ACG
rs13847	AAGCACACCGGCACGAAC	ACT
rs738658	GAGGCATTTTCATTATGCATG	CGT

[0210] The MassEXTEND™ reaction was performed in a total volume of 9 µl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0211] Following incubation, samples were desalted by adding 16 µl of water (total reaction volume was 25 µl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

#### Genetic Analysis

[0212] Minor allelic frequencies for the polymorphisms set forth in Table A were verified as being 10% or greater using the extension assay described above in a group of samples isolated from 92 individuals originating from the state of Utah in the United States, Venezuela and France (Coriell cell repositories).

[0213] Genotyping results are shown for female pools in Table 5. In Table 5, "AF" refers to allelic frequency; and "F case" and "F control" refer to female case and female control groups, respectively.

**TABLE 5: Genotyping Results**

SNP Reference	AF F case	AF F control	p-value
rs552	A = 0.190 G = 0.810	A = 0.123 G = 0.877	0.0011
rs12904	A = 0.455 G = 0.545	A = 0.375 G = 0.625	0.0012
rs2282146	C = 0.906 T = 0.094	C = 0.939 T = 0.061	0.0105
rs734784	G = 0.483 A = 0.517	G = 0.416 A = 0.584	0.0052
rs1042164	T = 0.233 C = 0.767	T = 0.159 C = 0.841	0.0002
rs749670	C = 0.342 T = 0.658	C = 0.419 T = 0.581	0.0038
rs955592	T = 0.045 C = 0.955	T = 0.076 C = 0.924	0.0177

SNP Reference	AF F case	AF F control	p-value
rs1143016	T = 0.093 C = 0.907	T = 0.054 C = 0.946	0.0071
rs755248	G = 0.146 A = 0.854	G = 0.069 A = 0.931	0.0000
rs1055055	A = 0.432 G = 0.568	A = 0.355 G = 0.645	0.0046
rs835409	T = 0.620 G = 0.380	T = 0.681 G = 0.319	0.0222
rs927663	T = 0.301 G = 0.699	T = 0.358 G = 0.642	0.0289
rs8162	A = 0.591 G = 0.409	A = 0.657 G = 0.343	0.0149
rs831038	C = 0.617 T = 0.383	C = 0.666 T = 0.334	0.0359
rs33079	G = 0.823 A = 0.177	G = 0.881 A = 0.119	0.0013
rs1710880	C = 0.303 A = 0.697	C = 0.371 A = 0.629	0.0129
rs1078153	T = 0.818 A = 0.182	T = 0.875 A = 0.125	0.0039
rs799570	A = 0.675 G = 0.325	A = 0.740 G = 0.260	0.0100
rs1282730	G = 0.086 A = 0.914	G = 0.127 A = 0.873	0.0150
rs1518875	T = 0.033 C = 0.967	T = 0.055 C = 0.945	0.0508
rs1568694	G = 0.045 A = 0.955	G = 0.081 A = 0.919	0.0084
rs905042	A = 0.832 T = 0.168	A = 0.769 T = 0.231	0.0047
rs1957723	G = 0.778 A = 0.222	G = 0.839 A = 0.161	0.0048
rs794018	G = 0.273 A = 0.727	G = 0.220 A = 0.780	0.0034
rs707723	C = 0.759 T = 0.241	C = 0.811 T = 0.189	0.0195
rs893861	G = 0.246 A = 0.754	G = 0.196 A = 0.804	0.0251
rs1914903	G = 0.861 A = 0.139	G = 0.910 A = 0.090	0.0055
rs2062232	C = 0.064 T = 0.936	C = 0.117 T = 0.883	0.0012
rs26609	A = 0.777 T = 0.223	A = 0.840 T = 0.160	0.0039
rs1370987	A = 0.422 G = 0.578	A = 0.341 G = 0.659	0.0007
rs1012414	G = 0.876 A = 0.124	G = 0.833 A = 0.167	0.0289
rs435903	G = 0.766 A = 0.234	G = 0.685 A = 0.315	0.0013
rs1248	T = 0.668 A = 0.332	T = 0.593 A = 0.407	0.0014
rs703508	G = 0.875 A = 0.125	G = 0.910 A = 0.090	0.0375
rs226465	G = 0.094 C = 0.906	G = 0.129 C = 0.871	0.0454
rs241448	C = 0.294 T = 0.706	C = 0.212 T = 0.788	0.0010
rs763155	A = 0.160	A = 0.114	0.0140



SNP Reference	AF F case	AF F control	p-value
	C = 0.840	C = 0.886	
	T = 0.069	T = 0.098	
rs1040461	C = 0.931	C = 0.902	0.0281
	A = 0.218	A = 0.145	
rs462832	T = 0.782	T = 0.855	0.0008
	T = 0.583	T = 0.679	
rs804194	C = 0.417	C = 0.321	0.0004
	A = 0.169	A = 0.103	
rs1022646	G = 0.831	G = 0.897	0.0007
	G = 0.853	G = 0.812	
rs1569112	A = 0.147	A = 0.188	0.0468
	A = 0.097	A = 0.140	
rs805623	G = 0.903	G = 0.860	0.0143
	A = 0.330	A = 0.240	
rs1019850	T = 0.670	T = 0.760	0.0005
	A = 0.581	A = 0.659	
rs1599931	G = 0.419	G = 0.341	0.0037
	A = 0.506	A = 0.577	
AA	G = 0.494	G = 0.423	0.0102
	T = 0.100	T = 0.138	
rs279941	G = 0.900	G = 0.862	0.0324
	C = 0.778	C = 0.717	
rs1062230	T = 0.222	T = 0.283	0.0109
	T = 0.295	T = 0.243	
rs1859911	C = 0.705	C = 0.757	0.0328
	T = 0.861	T = 0.809	
rs1477261	A = 0.139	A = 0.191	0.0105
	G = 0.121	G = 0.078	
rs1191119	A = 0.879	A = 0.922	0.0079
	A = 0.674	A = 0.583	
rs657780	G = 0.326	G = 0.417	0.0009
	G = 0.639	G = 0.724	
rs1393890	C = 0.361	C = 0.276	0.0014
	G = 0.331	G = 0.269	
rs1478714	A = 0.669	A = 0.731	0.0136
	C = 0.078	C = 0.044	
rs868213	T = 0.922	T = 0.956	0.0083
	G = 0.839	G = 0.784	
rs690115	A = 0.161	A = 0.216	0.0111
	A = 0.846	A = 0.903	
rs1465501	G = 0.154	G = 0.097	0.0020
	C = 0.895	C = 0.858	
rs899173	T = 0.105	T = 0.142	0.0408
	C = 0.087	C = 0.146	
rs10477	T = 0.913	T = 0.854	0.0010
	C = 0.715	C = 0.647	
rs926393	T = 0.285	T = 0.353	0.0082
	C = 0.194	C = 0.130	
rs465271	T = 0.806	T = 0.870	0.0019
	A = 0.111	A = 0.163	
rs13847	G = 0.889	G = 0.837	0.0056
	C = 0.898	C = 0.855	
rs738658	A = 0.102	A = 0.145	0.0183

[0214] All of the single marker alleles set forth in Table A were considered validated, since the genotyping data agreed with the allelotyping data and each SNP significantly associated with

osteoarthritis. Particularly significant associations with osteoarthritis are indicated by a calculated p-value of less than 0.05 for genotypic results.

### Example 3

#### Association of Polymorphic Variants with Osteoarthritis in Replication Cohorts

[0215] The single marker polymorphisms set forth in Table A were genotyped again in two replication cohorts consisting of individuals selected for OA.

#### Sample Selection and Pooling Strategies – Replication Sample 1

[0216] A second case control sample (replication sample #1) was created by using 100 Caucasian female cases from Chingford, UK, and 148 unrelated female cases from the St. Thomas twin study. Cases were defined as having Kellgren-Lawrence (KL) scores of at least 2 in at least one knee x-ray. In addition, 199 male knee replacement cases from Nottingham were included. (For a cohort description, see the Nottingham description provided in Example 1). The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The replication sample 1 cohort was used to replicate the initial results. Table 6 below summarizes the selected phenotype data collected from the case and control individuals.

**TABLE 6**

Phenotype	Female cases (n=248): median (range)/ (n,%)	Male cases (n=199): median (range)/ (n,%)	Female controls (n=313): mean (range)/ (n,%)
Age	59 (39- 73)	66 (45- 73)	55 (50- 72)
Height (cm)	162 (141- 178)	175 (152- 198)	162 (141- 176)
Weight (kg)	68 (51- 123)	86 (62- 127)	64 (40- 111)
Body mass index (kg/m <sup>2</sup> )	26 (18- 44)	29 (21- 41)	24 (18- 46)
Kellgren- Lawrence* left knee	0 (63, 26%), 1 (20, 8%), 2 (105, 43%), 3 (58, 23%), 4 (1, 0%)	NA	NA
Kellgren- Lawrence* right knee	0 (43, 7%), 1 (18, 7%), 2 (127, 52%), 3 (57, 23%), 4 (1, 0%)	NA	NA
KL* >2 both knees	No (145, 59%), Yes (101, 41%)	NA	NA

Phenotype	Female cases (n=248): median (range)/ (n,%)	Male cases (n=199): median (range)/ (n,%)	Female controls (n=313): mean (range)/ (n,%)
KL* >2 either knee	No (0, 0%), Yes (248, 100%)	NA	NA

\* 0: normal, 1: doubtful, 2: definite osteophyte (bony protuberance), 3: joint space narrowing (with or without osteophyte), 4: joint deformity

#### Sample Selection and Pooling Strategies – Replication Sample 2

[0217] A third case control sample (replication sample #2) was created by using individuals with symptoms of OA from Newfoundland, Canada. These individuals were recruited and examined by rheumatologists. Affected joints were x-rayed and a final diagnosis of definite or probable OA was made according to American College of Rheumatology criteria by a single rheumatologist to avoid any inter-examiner diagnosis variability. Controls were recruited from volunteers without any symptoms from the musculoskeletal system based on a normal joint exam performed by a rheumatologist. Only cases with a diagnosis of definite OA were included in the study. Only individuals of Caucasian origin were included. The cases consisted of 228 individuals with definite knee OA, 106 individuals with definite hip OA, and 74 individuals with hip OA.

**TABLE 7**

Phenotype	Case	Control
Age at Visit	62.7	52.5
Sex (Female/Male)	227/119	174/101
Knee OA Xray: No	35% (120)	80% (16)
Unknown	1% (4)	0% (0)
Yes	64% (221)	20% (4)
Hip OA Xray: No	63% (215)	80% (16)
Unknown	2% (7)	0% (0)
Yes	35% (121)	20% (4)

#### Assay for Verifying, Allelotyping, and Genotyping SNPs

[0218] Genotyping of the replication cohorts described in Tables 6 and 7 was performed using the same methods used for the original genotyping, as described herein. A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic

target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0219] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which were used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension probes used for analyzing (e.g., genotyping) polymorphisms in the replication cohorts. The initial PCR amplification reaction was performed in a 5 µl total volume containing 1X PCR buffer with 1.5 mM MgCl<sub>2</sub> (Qiagen), 200 µM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

[0220] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 µl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 µl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0221] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 7, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

[0222] The MassEXTEND™ reaction was performed in a total volume of 9 µl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0223] Following incubation, samples were desalted by adding 16 µl of water (total reaction volume was 25 µl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

#### Genetic Analysis

[0224] Genotyping results for replication cohorts #1 and #2 are provided in Tables 8 and 9, respectively.

TABLE 8

rsID	Replication #1 (Mixed Male/Female cases and Female controls)				Meta-analysis Disc. + Rep #1
	AF OA Con	AF OA Cas	Delta	P-value	P-value
rs552	0.87	0.85	0.02	0.344	0.0300
rs12904	0.57	0.57	0.00	0.936	0.2700
rs2282146	0.08	0.1	-0.02	0.342	0.0190
rs734784	0.52	0.54	-0.02	0.451	0.7200
rs1042164	0.79	0.82	-0.03	0.161	0.9100
rs749670	0.62	0.66	-0.04	0.173	0.0019
rs955592	0.93	0.94	-0.01	0.521	0.0600
rs1143016	0.93	0.93	0.00	0.869	NA
rs755248	0.9	0.89	0.01	0.544	0.1600
rs1055055	0.64	0.64	0.00	0.947	0.3300
rs835409	0.34	0.35	-0.01	0.715	0.1300
rs927663	0.64	0.65	-0.01	0.611	0.0690
rs831038	0.35	0.37	-0.02	0.399	NA
rs33079	0.14	0.14	0.00	0.995	0.3100
rs1710880	0.66	0.62	0.04	0.087	0.9000
rs799570	0.29	0.29	0.00	0.903	0.2500
rs1282730	0.88	0.87	0.01	0.751	0.4800
rs1568694	0.93	0.94	0.00	0.928	0.2600
rs905042	0.21	0.2	0.01	0.829	0.2200
rs1957723	0.13	0.16	-0.03	0.124	0.0009
rs794018	0.74	0.72	0.02	0.518	0.0710
rs707723	0.18	0.19	-0.01	0.658	0.0650
rs1914903	0.15	0.14	0.01	0.605	0.5500
rs2062232	0.91	0.91	0.00	0.788	0.2100
rs26609	0.16	0.19	-0.02	0.226	0.0032

rsID	Replication #1 (Mixed Male/Female cases and Female controls)				Meta-analysis Disc. + Rep #1
	AF OA Con	AF OA Cas	Delta	P-value	
rs1370987	0.63	0.63	-0.01	0.857	0.3900
rs1012414	0.12	0.13	-0.01	0.669	0.5600
rs435903	0.27	0.27	0.00	0.950	0.2800
rs1248	0.36	0.36	0.00	0.917	0.2400
rs703508	0.11	0.12	-0.01	0.558	0.0660
rs226465	0.87	0.88	-0.01	0.436	0.0500
rs241448	0.74	0.75	-0.01	0.805	0.4100
rs763155	0.86	0.88	-0.02	0.273	0.8800
rs1040461	0.92	0.92	0.00	0.826	NA
rs1022646	0.85	0.87	-0.02	0.219	0.8200
rs1569112	0.16	0.18	-0.02	0.402	0.8800
rs805623	0.87	0.88	-0.01	0.460	0.0370
rs1019850	0.69	0.7	0.00	0.890	0.3700
AA	0.47	0.48	-0.01	0.681	0.1200
rs279941	0.87	0.89	-0.01	0.400	0.0340
rs1062230	0.26	0.26	0.00	0.896	0.4200
rs1859911	0.71	0.75	-0.04	0.128	0.9000
rs1477261	0.16	0.16	0.00	0.986	0.3000
rs1191119	0.89	0.88	0.01	0.569	0.1200
rs1393890	0.29	0.31	-0.02	0.527	0.1400
rs1478714	0.69	0.67	0.03	0.300	0.0140
rs868213	0.92	0.93	-0.01	0.455	0.7000
rs890115	0.2	0.21	-0.01	0.729	0.4900
rs1465501	0.11	0.1	0.01	0.718	0.5600
rs899173	0.1	0.11	0.00	0.924	0.3300
rs10477	0.89	0.88	0.01	0.691	0.4700
rs926393	0.3	0.31	-0.01	0.830	0.4200
rs465271	0.86	0.85	0.01	0.516	0.0660
rs13847	0.86	0.85	0.01	0.547	0.5900
rs738658	0.14	0.15	-0.01	0.536	0.6700

TABLE 9

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis Disc. + Rep #2 Not Done
	AF OA Con	AF OA Cas	Delta	P-value	
rs552	0.85	0.86	-0.014	0.496	
rs12904	0.58	0.57	0.011	0.719	
rs2282146	0.08	0.08	0.002	0.876	
rs734784	0.53	0.54	-0.003	0.907	
rs1042164	0.83	0.80	0.026	0.248	
rs749670	0.66	0.62	0.036	0.208	
rs955592	0.95	0.92	0.033	0.027	

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis Disc. + Rep #2
	AF OA Con	AF OA Cas	Delta	P-value	Not Done
rs1143016	0.96	0.94	0.015	0.236	
rs755248	0.89	0.90	-0.009	0.608	
rs1055055	0.64	0.61	0.034	0.249	
rs835409	0.36	0.31	0.047	0.101	
rs927663	0.67	0.68	-0.013	0.631	
rs831038	0.34	0.35	-0.014	0.612	
rs33079	0.17	0.19	-0.019	0.417	
rs1710880	0.64	0.62	0.029	0.309	
rs799570	0.35	0.30	0.058	0.033	
rs1282730	0.89	0.89	-0.001	0.982	
rs1568694	0.95	0.94	0.009	0.518	
rs905042	0.19	0.20	-0.002	0.933	
rs1957723	0.18	0.20	-0.017	0.454	
rs794018	0.73	0.72	0.015	0.586	
rs707723	0.20	0.21	-0.007	0.759	
rs1914903	0.14	0.16	-0.022	0.285	
rs2062232	0.92	0.91	0.008	0.632	
rs26609	0.19	0.18	0.005	0.827	
rs1370987	0.59	0.61	-0.023	0.423	
rs1012414	0.15	0.14	0.008	0.679	
rs435903	0.24	0.26	-0.026	0.316	
rs1248	0.33	0.38	-0.051	0.078	
rs703508	0.10	0.11	-0.002	0.916	
rs226465	0.89	0.89	-0.007	0.699	
rs241448	0.76	0.77	-0.007	0.778	
rs763155	0.89	0.84	0.049	0.016	
rs1040461	0.91	0.91	0.001	0.948	
rs1022646	0.86	0.86	-0.001	0.974	
rs1569112	0.16	0.17	-0.016	0.446	
rs805623	0.89	0.87	0.022	0.256	
rs1019850	0.71	0.69	0.026	0.341	
AA	0.48	0.44	0.035	0.234	
rs279941	0.91	0.87	0.037	0.047	
rs1062230	0.23	0.22	0.011	0.653	
rs1859911	0.72	0.71	0.015	0.560	
rs1477261	0.17	0.14	0.031	0.143	
rs1191119	0.86	0.88	-0.017	0.377	
rs1393890	0.30	0.28	0.017	0.516	
rs1478714	0.68	0.70	-0.025	0.358	
rs868213	0.91	0.93	-0.019	0.260	
rs690115	0.19	0.18	0.005	0.811	
rs1465501	0.10	0.12	-0.020	0.282	
rs899173	0.14	0.12	0.020	0.319	
rs10477	0.86	0.88	-0.016	0.442	
rs926393	0.37	0.32	0.042	0.137	

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis Disc. + Rep #2
	AF OA Con	AF OA Cas	Delta	P-value	Not Done
rs465271	0.87	0.85	0.023	0.263	
rs13847	0.84	0.85	-0.012	0.582	
rs738658	0.18	0.15	0.021	0.340	

[0225] To combine the evidence for association from multiple sample collections, a meta-analysis procedure was employed. The allele frequencies were compared between cases and controls within the discovery sample, as well as within the replication cohort #1 using the DerSimonian-Laird approach (DerSimonian, R. and N. Laird. 1986. Meta-analysis in clinical trials. Control Clin Trials 7: 177-188.)

[0226] The absence of a statistically significant association in one or more of the replication cohorts should not be interpreted as minimizing the value of the original finding. There are many reasons why a biologically derived association identified in a sample from one population would not replicate in a sample from another population. The most important reason is differences in population history. Due to bottlenecks and founder effects, there may be common disease predisposing alleles present in one population that are relatively rare in another, leading to a lack of association in the candidate region. Also, because common diseases such as arthritis-related disorders are the result of susceptibilities in many genes and many environmental risk factors, differences in population-specific genetic and environmental backgrounds could mask the effects of a biologically relevant allele. For these and other reasons, statistically strong results in the original, discovery sample that did not replicate in one or more of the replication samples may be further evaluated in additional replication cohorts and experimental systems.

#### Example 4

##### KIAA0296 Region Proximal SNPs

[0227] SNP rs749670 is associated with osteoarthritis and is described in Table A. It lies within the KIAA0296 gene and codes for a G327E amino acid change. The thymine allele of SNP rs749670 is associated with osteoarthritis (see Table 5) and codes for glutamic acid. KIAA0296 shares homology with C2H2-type Zn-finger protein and is likely a novel transcription factor. One-hundred one additional allelic variants proximal to rs749670 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 10. The chromosome positions provided in column four of Table 10 are based on Genome "Build 34" of NCBI's GenBank.



TABLE 10

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs7500176	16	247	31077197	a/g
rs6565212	16	1535	31078485	c/t
rs8054046	16	2386	31079336	c/t
rs8056842	16	6440	31083390	c/t
rs732173	16	9133	31086083	g/t
rs732172	16	9143	31086093	a/g
rs7188557	16	9471	31086421	a/t
rs2288004	16	13150	31090100	c/g
rs4337310	16	13717	31090667	c/t
rs2016554	16	14466	31091416	a/g
rs6565213	16	15769	31092719	a/c
rs7204762	16	16870	31093820	a/g
rs4889529	16	18545	31095495	c/t
rs6565214	16	18749	31095699	c/t
rs7499674	16	19123	31096073	g/t
rs6565215	16	20736	31097686	a/g
rs1023623	16	21038	31097988	c/t
rs1023624	16	21046	31097996	c/t
rs1023625	16	21050	31098000	c/t
rs1549297	16	21056	31098006	a/t
rs3084894	16	21706	31098656	-/acc
rs8048228	16	23170	31100120	a/g
rs7405432	16	25028	31101978	a/t
rs8054249	16	27871	31104821	a/g
rs8061047	16	28070	31105020	c/t
rs7187220	16	31717	31108667	a/g
rs8046978	16	32019	31108969	a/g
rs2288003	16	32318	31109268	a/g
rs7196421	16	33080	31110030	a/g
rs7196431	16	33101	31110051	a/g
rs7203158	16	34236	31111186	a/g
rs2303223	16	34285	31111235	c/t
rs2032917	16	34818	31111768	c/g
rs8044134	16	35168	31112118	c/g
rs4889531	16	37981	31114931	c/t
rs4889532	16	38113	31115063	c/g
rs4889533	16	38117	31115067	c/t
rs881929	16	38481	31115431	g/t
rs8047104	16	38615	31115565	c/g
rs8047803	16	38944	31115894	a/c
rs4644874	16	39288	31116238	a/c
rs2359673	16	41385	31118335	c/t
rs4435271	16	42136	31119086	a/t
rs7197717	16	42185	31119135	a/c
rs2359674	16	42353	31119303	a/g
rs6565217	16	42434	31119384	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs2303222	16	44580	31121530	a/g
rs4889615	16	44675	31121625	a/t
rs4624197	16	45739	31122689	g/t
rs3751853	16	46439	31123389	c/t
rs749671	16	47457	31124407	c/t
rs749670	16	47735	31124685	c/t
rs3751855	16	50319	31127269	c/t
rs3751856	16	50708	31127658	a/g
rs7196726	16	51185	31128135	a/g
rs889550	16	53002	31129952	a/g
rs750952	16	53064	31130014	c/t
rs2077633	16	53637	31130587	a/g
rs7199949	16	55274	31132224	c/g
rs2032916	16	55825	31132775	c/t
rs4468641	16	55986	31132936	a/c
rs4889535	16	56684	31133634	c/g
rs4316775	16	57653	31134603	c/t
rs4313819	16	57659	31134609	c/g
rs6565218	16	57692	31134642	g/t
rs4318224	16	57775	31134725	c/t
rs1046030	16	61313	31138263	c/t
rs7294	16	61431	31138381	a/g
rs7200749	16	61699	31138649	a/g
rs2359612	16	62906	31139856	a/g
rs8050894	16	63619	31140569	c/g
rs2884737	16	64664	31141614	a/c
rs1895514	16	68452	31145402	g/t
rs8060209	16	69665	31146615	c/t
rs8060217	16	69681	31146631	c/t
rs7196161	16	70091	31147041	a/g
rs8062336	16	74637	31151587	a/g
rs8043778	16	74760	31151710	a/g
rs2032915	16	76523	31153473	a/g
rs4889616	16	78559	31155509	c/g
rs1045564	16	79549	31156499	a/c
rs2303221	16	79882	31156832	c/t
rs1549296	16	81339	31158289	a/g
rs889555	16	81681	31158631	c/t
rs5816521	16	81696	31158646	-/g
rs749767	16	83517	31160467	c/t
rs2884738	16	85431	31162381	a/c
rs2052581	16	86332	31163282	c/t
rs4889617	16	87358	31164308	a/g
rs4889619	16	87725	31164675	c/t
rs1978487	16	89052	31166002	a/g
rs1978486	16	90020	31166970	a/g
rs1978485	16	90231	31167181	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs4889620	16	90284	31167234	a/g
rs4889621	16	90447	31167397	c/t
rs3214477	16	90601	31167551	-/g
rs4527034	16	90724	31167674	a/g
rs1060506	16	92559	31169509	c/t
rs7200125	16	95176	31172126	a/g
rs6565219	16	95195	31172145	c/t
rs889548	16	96822	31173772	a/g

# Assay for Verifying and Allelotyping SNPs

[0228] The methods used to verify and allelotype the 101 proximal SNPs of Table 10 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 11 and Table 12, respectively.

TABLE 11

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs7500176	ACGTTGGATGACAGTGGCTCATGCCGTGTA	ACGTTGGATGTTTACCATAATTGGCCAGGC
rs6565212	ACGTTGGATGTTAGGAAGGATGTGGAAGGG	ACGTTGGATGGACCTGACCTCAAGAGAGAAG
rs8054046	ACGTTGGATGCATGGAAGTTTAGAGCAGCC	ACGTTGGATGTGCACAGTGGGTAACCTGTAG
rs8056842	ACGTTGGATGATGAGGTTTCACCTTGTTGG	ACGTTGGATGATCATAGCACTTTGCGAGGC
rs732173	ACGTTGGATGAGACCAGGCTCAGTCCAAAC	ACGTTGGATGTGGCCAAACCTGGAAGACAC
rs732172	ACGTTGGATGCTCAGTCCAAACTGCCAGAC	ACGTTGGATGCATGGCCAAACCTGGAAGAC
rs7188557	ACGTTGGATGAACATCTGTACAAGGCTGGG	ACGTTGGATGATTGGCTGTAGCATGACTGA
rs2288004	ACGTTGGATGAAGACACTGGAAGGCTGTG	ACGTTGGATGAGAGAAGGTGGAGCTCTTTC
rs4337310	ACGTTGGATGAGGGAAGAGATGTACACAGG	ACGTTGGATGTTTGGAGCAGACTGGTAGG
rs2016554	ACGTTGGATGAAGCAATCCTCCACCTCAG	ACGTTGGATGCAAGAGCAAAACTCCCTCTC
rs6565213	ACGTTGGATGAGATGGAGTCTCACTCCATC	ACGTTGGATGTGAGGCGAGGAGAATCGCTTG
rs7204762	ACGTTGGATGAGTGGCTCACACCTGTAATC	ACGTTGGATGGCTGGTCTTGAACCTCTGAC
rs4889529	ACGTTGGATGCAAGCAATCCTTGCCTCAAG	ACGTTGGATGGTGGTTCACATCTGCAATC
rs6565214	ACGTTGGATGTGATCTCGGCTCACTGCAAG	ACGTTGGATGAAAATTAGCCGGCATGGTG
rs7499674	ACGTTGGATGAAGTACGGGAACCTTCCCAC	ACGTTGGATGTGGGCCCCCACTAAGCTTAA
rs6565215	ACGTTGGATGAGACGGAAGTCCAGCTTTG	ACGTTGGATGTGGGAACCACTCTGTTCTATG
rs1023623	ACGTTGGATGACAGGCAAGACTCCATCTC	ACGTTGGATGTCCTCTTACAGGCTGTTTAC
rs1023624	ACGTTGGATGTGACAGAGCAAGACTCCATC	ACGTTGGATGGTCTTAACCACTGAGCCTAT
rs1023625	ACGTTGGATGTGGTGACAGAGCAAGACTCC	ACGTTGGATGTCAGGTCCTAACCAGTGAGC
rs1549297	ACGTTGGATGTTGCATTGATCCGAGATCGC	ACGTTGGATGTCAGGTCCTAACCAGTGAGC
rs3084894	ACGTTGGATGTCCAGGTTCAAGCGATTCT	ACGTTGGATGCCATGAAACCCCATCTCTAC
rs8048228	ACGTTGGATGAATTGCTGAACCTTGGGAGG	ACGTTGGATGTTTGACAGTCTCCCTCTATC
rs7405432	ACGTTGGATGAGATCATGCCACTGCACTAC	ACGTTGGATGCATGCACTTGGCCCTAATTG
rs8054249	ACGTTGGATGATCTCCTGACCTCATGATCT	ACGTTGGATGTAATCAACACCAAGGCTGGG
rs8061047	ACGTTGGATGATGATCACAGCTCACTGACG	ACGTTGGATGCTCCCTGCGCTCTACAAAAAG
rs7187220	ACGTTGGATGAAGGAGACCTTCTCCACAA	ACGTTGGATGCGCTGTCAGAGAGGCTCTTG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs8046978	ACGTTGGATGTGCACAGGAGCTGGTGGTG	ACGTTGGATGATCACACCACTGACTCCGG
rs2288003	ACGTTGGATGACCGGCCGTTCAAGTGCCTG	ACGTTGGATGAGAGTGCACCAAGCGCGTGC
rs7196421	ACGTTGGATGTTCAACGCCATTCTCCTGCCT	ACGTTGGATGAAATTAGCCAGGCGTGGTGG
rs7196431	ACGTTGGATGAGATCTCGGCTCACTGCAAG	ACGTTGGATGATGTAGTCCAGCTACTCGG
rs7203158	ACGTTGGATGAAGCCTATGCGGAGCTCAAG	ACGTTGGATGATTGGCTGCAGCAACGCTGT
rs2303223	ACGTTGGATGACCCCTACCGCTCATGGTTG	ACGTTGGATGTGCGGCCCTACAGCTGTGA
rs2032917	ACGTTGGATGCCTGGGCGCGTTTGGAAATG	ACGTTGGATGAGCCCCCGGCTACAAGCGCT
rs8044134	ACGTTGGATGACTAAGAAAGGAGGCTGAGG	ACGTTGGATGACAGTGTITTTGAAAAGCCCG
rs4889531	ACGTTGGATGATTCCTCACCCAACCTCTGTC	ACGTTGGATGGACCGTGTGTAATGTACTGCG
rs4889532	ACGTTGGATGGGGACAAGAATCCCTATCTC	ACGTTGGATGTAGAGCCACACACATTGCTG
rs4889533	ACGTTGGATGCTCTGTAAGATAGGGACAAG	ACGTTGGATGTAGAGCCACACACATTGCTG
rs861929	ACGTTGGATGTTGACCCAGTGGTCTGAGC	ACGTTGGATGCCAGCTACCTGGTGTCTAAC
rs8047104	ACGTTGGATGGTGGGATGTTAGACAGAGAC	ACGTTGGATGTGCCAGGTTGGTCTCAGCAT
rs8047803	ACGTTGGATGAAAGTCTGGGATTACAGCG	ACGTTGGATGAAATACAGGATTCCTGAGGCC
rs4644874	ACGTTGGATGAGCTTGCTATGTTGCCTGG	ACGTTGGATGTAATCCAGCACTTTGGGAG
rs2359673	ACGTTGGATGGGTGTGATGTCAGTTCACCTGC	ACGTTGGATGATCCCAAACTTTGGGAGGC
rs4435271	ACGTTGGATGACAGCTGGTCTCAAGAACTCC	ACGTTGGATGTGGCTCATGCCCTGTAATCAC
rs7197717	ACGTTGGATGTTGGGAAAAGTCTCCAGAG	ACGTTGGATGGCTGCAAGGAGTATGCTCC
rs2359674	ACGTTGGATGGCCTAGCAGTTCATTATGAG	ACGTTGGATGCCTTGTCTCCAAATACAGTC
rs6565217	ACGTTGGATGAAGAACGCTAATCCTACTGG	ACGTTGGATGTGGAGACAAGGCCTTTATGG
rs2303222	ACGTTGGATGTTGGGAAAAGTCTCCAGAG	ACGTTGGATGGCGCAGAAAGGGAGAAAAAG
rs4889615	ACGTTGGATGTAAGTCTTAGGCTGCACGG	ACGTTGGATGATGCACCGGAACGATTCTAG
rs4624197	ACGTTGGATGCGCTAAGAGAGTCTTTTGGG	ACGTTGGATGCAGAGCGAGACTCCATCTCA
rs3751853	ACGTTGGATGTCCCTTAGGCTTAAGTCATC	ACGTTGGATGGGTCTGTGATCAGAAAGTAGG
rs749671	ACGTTGGATGTGACTACATTTGTACCGCCG	ACGTTGGATGTCAAGTAGTGAACCTTCACAGG
rs749670	ACGTTGGATGTCTCATCTGTGTGCCATTG	ACGTTGGATGATGAGGGTGAAGGCGAGGAG
rs3751855	ACGTTGGATGAAGAAGAGGTGTGGGAGGAG	ACGTTGGATGTCAAGCTGGCTTCAGTCTG
rs3751856	ACGTTGGATGAGCTGTACTGGCCCGTCTCG	ACGTTGGATGCAAGTGCAGGCGGACCTATC
rs7196726	ACGTTGGATGGACCTAGTTAGGAACTAGG	ACGTTGGATGTGAGGCGAGCAAGCTGAGAG
rs889550	ACGTTGGATGTCCACCAGCACTGCTGGA	ACGTTGGATGCAGGTCCTGCTGAGGGAAAC
rs750952	ACGTTGGATGTTCCCTCAGCAGGACCTGG	ACGTTGGATGGGTGGCCACTAGATGGAATG
rs2077633	ACGTTGGATGTTTCTCAGGAGTAGTTCCGG	ACGTTGGATGAAAGCAAGCCAGATCTGGGTC
rs7199949	ACGTTGGATGTCCCATCAGGCAGGTGGT	ACGTTGGATGCAGCCTGTGACACTGGGAG
rs2032916	ACGTTGGATGGTTCCTCCTACTACTGAAGG	ACGTTGGATGTGCCACTTGGCTGTAGTTAC
rs4468641	ACGTTGGATGATGAGTACAGGAATACGGGAG	ACGTTGGATGAAATGCCCTACTTGTCACTC
rs4889535	ACGTTGGATGCTATGGCAGACACCCCTCTGA	ACGTTGGATGGAAGAGAAGGAGCAGAAAGG
rs4316775	ACGTTGGATGAGTAGCTCACGCTTGTAAATC	ACGTTGGATGCTATGTTGACACAGGCTAGTC
rs4313819	ACGTTGGATGTGCTCAGAGGCTAGTCTTGAAC	ACGTTGGATGAGTAGCTACAGCTTGTAAATC
rs5655218	ACGTTGGATGTTAAAGTCAACAGACTGAGGC	ACGTTGGATGTTGAACCTTTGGGCTCAAGC
rs4318224	ACGTTGGATGTCAGTCTGTGACTTTAAGCG	ACGTTGGATGACCACCTTTTATGGTAGAAG
rs1046030	ACGTTGGATGGTCTCCAAAGCTCTTTCCATT	ACGTTGGATGGGATGTAGTCAAGAACTTTA
rs7294	ACGTTGGATGGCACTGGGTGTAATAAAGAG	ACGTTGGATGTTCTAGATTACCCCTCCTC
rs7200749	ACGTTGGATGGAGCAGCAAGAACAGGATCC	ACGTTGGATGCTGCTGCTGAGTCTGCTGAG
rs2359612	ACGTTGGATGAAATCGGCCAAGTCTGAACC	ACGTTGGATGTCCAGAGCAAGGCACTAGTA
rs8050894	ACGTTGGATGAATCTTGGTATGCACACAG	ACGTTGGATGTAGTTACCTCCCAACCTCCC
rs2884737	ACGTTGGATGTCTAATGCTAACGCTGGC	ACGTTGGATGTTGACGATGGTCTCAAGGAC
rs1895514	ACGTTGGATGCAATCTCAGCTCACTGCAAC	ACGTTGGATGTAATCCAGCTACTTGGGAG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs8060209	ACGTTGGATGGGTCAGGAGTTTAAAGACAAG	ACGTTGGATGCCATGCCCGGCTAATTTTGTG
rs8060217	ACGTTGGATGTGAGTAGCTGGGATTACAGG	ACGTTGGATGAGACAAGCTTGGCCAACATG
rs7196161	ACGTTGGATGGTGTCTTTTAGTAGAGACGGG	ACGTTGGATGATCCCAGCAGCTTTAGGAAGC
rs8062336	ACGTTGGATGTGCTCCCACTCTCAGACG	ACGTTGGATGAAGCGAGCGCGCTCTTC
rs8043778	ACGTTGGATGTCTCTCACTTCTCAGACGGG	ACGTTGGATGATCGTCTGAGATGTGGGGAG
rs2032915	ACGTTGGATGATCCCAACCCGTTCTTTCCC	ACGTTGGATGTTCCCGCTCCCTTTTACCAC
rs4889616	ACGTTGGATGGAACCAAGAACTGGAAGGAG	ACGTTGGATGTGTAAAGCGCACAGATCACG
rs1045564	ACGTTGGATGTGTCAAGCATCCTCGACGCAC	ACGTTGGATGACCCAGGCGACCCAAAATGG
rs2303221	ACGTTGGATGAGAACCCCAACACTCTCCC	ACGTTGGATGAGCGGAGAAGGTGCGCAAG
rs1549296	ACGTTGGATGATGCTGCTGAACCTTCTAAC	ACGTTGGATGAGCAGGGTTTCTCAACCATG
rs889555	ACGTTGGATGAGACCAGTAGGTACAAGCAC	ACGTTGGATGTCAAGAATGCCATGAGGTGG
rs5816521	ACGTTGGATGATTGTGGCTCTATGCAGAGG	ACGTTGGATGTCAAGAATGCCATGAGGTGG
rs749767	ACGTTGGATGCTGATAGAAAGGACCAAGGA	ACGTTGGATGCTGGAGTTCTGATTCAAGGTG
rs2884738	ACGTTGGATGAGAACTGCTTGAACCCAGGA	ACGTTGGATGATGGAGTCTTGTGTGTCCGG
rs2052581	ACGTTGGATGTGGACATGCGGATATGGAG	ACGTTGGATGAGGGGTTCTGTGAGAGTCAG
rs4889617	ACGTTGGATGCAGAGCGAGACTCCATCTCA	ACGTTGGATGACACTCGCGCTGGCCCTAATG
rs4889619	ACGTTGGATGAAAATTAGATGGGCGTGGTG	ACGTTGGATGATCTCGGCTCACTGCCAACCT
rs1978487	ACGTTGGATGTCCCCTCTCTATGTTCCCTGC	ACGTTGGATGATGGAGGAAGACAGAGAGAG
rs1978486	ACGTTGGATGTACCTAGGGTACAGATTGTG	ACGTTGGATGGGGTATGTGGTAAATGAGC
rs1978485	ACGTTGGATGTCAAGCAATTTTCTGCTC	ACGTTGGATGCCATCTGTACCAAAAAGAGC
rs4889620	ACGTTGGATGTGGCAAAACCCATCTGTAC	ACGTTGGATGAGTAGTTGGGAATTACAGGTG
rs4889621	ACGTTGGATGTACTCAATCACTGCCACAAC	ACGTTGGATGGCCAGTTATTTTCTCATTCG
rs3214477	ACGTTGGATGACTCGAGACTGGATCACTTC	ACGTTGGATGCCCTTTTGTTCACGCCCTTACC
rs4527034	ACGTTGGATGAAGTATGGGCCATAAGAGTG	ACGTTGGATGTATGTACACTACGTGGGCTG
rs1060506	ACGTTGGATGATCAGGAGTGCAAACCAAG	ACGTTGGATGGGATGAAGCTGCAATAGCTG
rs7200125	ACGTTGGATGATTTTCCATTGCACTCCAG	ACGTTGGATGTACAGGCATGAGCCATAGCC
rs6565219	ACGTTGGATGCTTGGCTCTCAAAGTGCTG	ACGTTGGATGAGGGCGAGGCTCATTTCAA
rs889548	ACGTTGGATGCTGGCCAAGTCTAATACAG	ACGTTGGATGCCCAATTCCAGAGATGTGAC

TABLE 12

dbSNP rs#	Extend Primer	Term Mix
rs7500176	GATCACGAGGTCAGGAGTTC	ACT
rs6565212	GCTGGAAAACGTGTGAGGGT	ACT
rs8054046	TTTAGAGCAGCCGATACCCA	ACG
rs8056842	GCTGGTCTCGAACTCCTGA	ACG
rs732173	GCTCAGTCCAAACTGCCAG	CGT
rs732172	ACTGCCAGACTCCCGCCA	ACG
rs7188557	CCTGGCCCTGGTTGTGAGT	CGT
rs2288004	CGGCAGATCCAGTGTGTC	ACT
rs4337310	CACGGAATCTCCAGTGCAC	ACT
rs2016554	GGCAGCTACCACTGACATG	ACG
rs6565213	GCAGTGGCGCAATCTTGAC	ACT
rs7204762	CCCAGCACTTGGGAGGC	ACG

dbSNP rs#	Extend Primer	Term Mix
rs4889529	CTCAAGTGATCCTCCTGCCT	ACG
rs6565214	GAGTAGCTGGGACTACAGG	ACG
rs7499674	GTTCTTCTCAACATCTGCCCA	ACT
rs6565215	TTTCCTTCAGACAGGGCTCT	ACT
rs1023623	GACTCCATCTCAAAAAAAAAAAAA	ACT
rs1023624	GAGCAAGACTCCATCTCAAAAA	ACT
rs1023625	CAGAGCAAGACTCCATCTCA	ACT
rs1549297	GGTGACAGAGCAAGACTCC	CGT
rs3084894	CGAGTAGGTGGGACTACAG	ACT
rs8048228	TGAGCCGAGATGGCAACAC	ACG
rs7405432	CTACAGGCTAGGAGACAGAG	CGT
rs8054249	AAAGTGCTGGGATTACAGGC	ACT
rs8061047	CCTCCTGAGGAGCTGGTCT	ACT
rs7187220	GGCCCTTCCCCTGCACC	ACG
rs8046978	AGAGTTCAGCCGCCCGG	ACG
rs2288003	GTGACAAGACGTTCTGTGGC	ACT
rs7196421	CTCAGCCTCCCGAGTAGC	ACG
rs7196431	CGGGTTACGCCATTCTCC	ACG
rs7203158	CAACCATGAGCGGTGAGGG	ACG
rs2303223	TTGAGCTCCGCATAGGCTTT	ACT
rs2032917	TGGAATGTCTTGGTACAGGACA	ACT
rs8044134	CCTACACGTCCCCCCC	ACT
rs4889531	CAACTCTGTCAAGTAAGTACT	ACT
rs4889532	CAAGAATCCCTATCTCAGAAAG	ACT
rs4889533	GGACAAGAATCCCTATCTCAG	ACT
rs881929	CTGCCTCTTGCCAGCTCTG	ACT
rs8047104	CAGAGACCTAGCCTACCTG	ACT
rs8047803	TTACAGGCGAGAGCCACCA	CGT
rs4644874	GGGCTCAAGTGATCCTCCC	CGT
rs2359673	ACTGCGACCTCTGCCTCC	ACG
rs4435271	GCTTCAGATGCTCCTCACT	CGT
rs7197717	GCATGAGCCGTGACACAGC	CGT
rs2359674	GAATGTTTGTGTTCCCTGTCC	ACT
rs6565217	CCAGGGCCATACCTTATGA	ACG
rs2303222	AAAGTGTCACCAAGTAC	ACG
rs4889615	GCGGCGTCTTTGCACGCTA	CGT
rs4624197	AGAGAGTCTTTGGGGTTTTTT	ACT
rs3751853	CCTACAGGTATAGCTAAGGAA	ACT
rs749671	ATTTGTACCGCCGCTCCTC	ACG
rs749670	GGTGGTGGGCATCCCTTTC	ACG
rs3751855	AGAGCCCAGGCTGGAGAC	ACG
rs3751856	CCGTCTCGTGGCTGCGC	ACG
rs7196726	GTTAGGAAGTGAAGACCCAG	ACG
rs889550	AGCACTGCTGGAAGCCGC	ACT

dbSNP rs#	Extend Primer	Term Mix
rs750952	GCTGGCCTCTCCACCTCC	ACG
rs2077633	CCATATCTTCTCCTCTCCCC	ACG
rs7199949	CAGGCAGGTGGTGGTCAG	ACT
rs2032916	CCAAAGTTCCAGAGAGGTTAA	ACT
rs4468641	ATACGGGAGGCAGGCCCA	ACT
rs4889535	CAGACACCCCTCTGATTGCAG	ACT
rs4316775	GAGGATCGCTTGAGCCCAA	ACT
rs4313819	GCTAGTCTTGAACCTCTGGG	ACT
rs6565218	CTCACGCTTGTAAATCCAGC	CGT
rs4318224	TTCCCTTGCAACCTGAGTTT	ACG
rs1046030	GCCCAGGGAGGGAAGGTT	ACG
rs7294	TTGGTCCATTGTCAATGTG	ACG
rs7200749	GAAGAACAGGATCCAGGCCA	ACT
rs2359612	CCATGTGTGTCAGCCAGGACC	ACT
rs8050894	CCAGCTAGCTGCTCATCAC	ACT
rs2884737	TCGCCAACACCCCCCTTC	CGT
rs1895514	CCCCTCTCGGGTTCAAGC	CGT
rs8060209	TGGCCAACATGGCGAAACC	ACG
rs8060217	CCATGCCCGGCTAATTTTTGT	ACT
rs7196161	AACTCCTGACCTCATGATCC	ACT
rs8062336	TCACTTCTAGATGGGAAGG	ACG
rs8043778	CGCTCCTCACCTCCCAGA	ACG
rs2032915	TTCTTCCCAACGTCCTGGA	ACT
rs4889616	GAAGTGAAGGAGGACAAGA	ACT
rs1045564	GTCCCTGAAGTCGGAGAAG	CGT
rs2303221	CTCTCCCTCCCGCTACAT	ACG
rs1549296	TGCACGGGGCAGCCCCCT	ACT
rs889555	AGCACCCCGGTTCTGTCC	ACT
rs5816521	CCAGTAGGTACAAGCACC	ACT
rs749767	GACCAAGGATTGGGCAAG	ACT
rs2884738	CCAGGAGGTGGAGGTTGCA	ACT
rs2052581	GGATATGGAGGGCCGATTGT	ACT
rs4889617	GAGACTCCATCTCAAAAAAAAA	ACT
rs4889619	GCAGAGGAATCGCTTGAACC	ACG
rs1978487	GTTCTGCAACATTTTTTCTTA	ACG
rs1978486	GGGTACAGATTTGAAAAGTG	ACT
rs1978485	TTTTCTGCCTCAGCCTCC	ACG
rs4889620	ACCCATCTGTACCAAAAAGA	ACG
rs4889621	CTGTGAGGTGGATCAGTTG	ACT
rs3214477	GCAGAACTGTGATGGAAAAAG	ACT
rs4527034	CCAGGGCAGCCAACCTCCC	ACG
rs1060506	AAGTCTCAGACCCAGA	ACG
rs7200125	AGGCTCCATTTCAAAAAAAAA	ACT
rs6565219	AAAGTGCTGGGATTACAGGC	ACT

dbSNP rs#	Extend Primer	Term Mix
rs889548	AGTCCTAATACAGTGGATGTC	ACT

## Genetic Analysis

[0229] Allelotyping results from the discovery cohort are shown for cases and controls in Table 13. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs732173 has the following case and control allele frequencies: case A1 (G) = 0.55; case A2 (T) = 0.45; control A1 (G) = 0.58; and control A2 (T) = 0.42, where the nucleotide is provided in paranthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 13

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs7500176	247	31077197	A/G			
rs6565212	1535	31078485	C/T			
rs8054046	2386	31079336	C/T			
rs8056842	6440	31083390	C/T			
rs732173	9133	31086083	G/T	0.45	0.42	0.382
rs732172	9143	31086093	A/G			
rs7188557	9471	31086421	A/T			
rs2288004	13150	31090100	C/G	0.52	0.45	0.026
rs4337310	13717	31090667	C/T	0.18	untyped	
rs2016554	14466	31091416	A/G			
rs6565213	15769	31092719	A/C			
rs7204762	16870	31093820	A/G			
rs4889529	18545	31095495	C/T			
rs6565214	18749	31095699	C/T			
rs7499674	19123	31096073	G/T			
rs6565215	20736	31097686	A/G			
rs1023623	21038	31097988	C/T	0.02	untyped	
rs1023624	21046	31097996	C/T	0.16	0.11	0.035
rs1023625	21050	31098000	C/T	0.32	NA	
rs1549297	21056	31098006	A/T			
rs3084894	21706	31098656	-ACC			
rs8048228	23170	31100120	A/G	0.54	0.61	0.040
rs7405432	25028	31101978	A/T	0.35	0.43	0.025
rs8054249	27871	31104821	A/G			
rs8061047	28070	31105020	C/T	0.21	0.21	0.903
rs7187220	31717	31108667	A/G			
rs8046978	32019	31108969	A/G	0.34	0.28	0.083
rs2288003	32318	31109268	A/G			
rs7196421	33080	31110030	A/G			
rs7196431	33101	31110051	A/G			
rs7203158	34236	31111186	A/G			
rs2303223	34285	31111235	C/T	0.52	0.45	0.060
rs2032917	34818	31111768	C/T			



dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs8044134	35168	31112118	C/G	0.97	0.97	0.856
rs4889531	37981	31114931	C/T			
rs4889532	38113	31115063	C/G			
rs4889533	38117	31115067	C/T			
rs861929	38481	31115431	G/T	0.38	0.34	0.228
rs8047104	38615	31115565	C/G	0.60	0.65	0.117
rs8047803	38944	31115894	A/C	0.35	0.33	0.437
rs4644874	39288	31116238	A/C			
rs2359673	41385	31118335	C/T	0.18	0.20	0.563
rs4435271	42136	31119086	A/T			
rs7197717	42185	31119135	A/C			
rs2359674	42353	31119303	A/G	0.22	0.18	0.122
rs6565217	42434	31119384	A/G	0.35	0.33	0.608
rs2303222	44580	31121530	A/G	0.60	0.52	0.022
rs4889615	44675	31121625	A/T			
rs4624197	45739	31122689	G/T			
rs3751853	46439	31123389	C/T			
rs749671	47457	31124407	C/T	0.32	0.37	0.095
rs749670	47735	31124685	C/T			
rs3751855	50319	31127269	C/T	0.53	0.57	0.287
rs3751856	50708	31127658	A/G			
rs7196726	51185	31128135	A/G	0.41	0.37	0.258
rs889550	53002	31129952	A/G			
rs750952	53064	31130014	C/T	0.43	0.41	0.535
rs2077633	53637	31130587	A/G			
rs7199949	55274	31132224	C/G	0.46	0.53	0.051
rs2032916	55825	31132775	C/T			
rs4468641	55986	31132936	A/C	0.26	0.25	0.902
rs4889535	56684	31133634	C/G			
rs4316775	57653	31134603	C/T			
rs4313819	57659	31134609	C/G			
rs6565218	57692	31134642	G/T			
rs4318224	57775	31134725	C/T			
rs1046030	61313	31138263	C/T			
rs7294	61431	31138381	A/G	0.38	0.37	0.669
rs7200749	61699	31138649	A/G			
rs2359612	62906	31139856	A/G	0.56	0.48	0.017
rs8050894	63619	31140569	C/G	0.48	0.45	0.320
rs2884737	64664	31141614	A/C	0.68	0.60	0.016
rs1895514	68452	31145402	G/T			
rs8060209	69665	31146615	C/T			
rs8060217	69681	31146631	C/T			
rs7196161	70091	31147041	A/G			
rs8062336	74637	31151587	A/G			
rs8043778	74760	31151710	A/G			
rs2032915	76523	31153473	A/G	0.43	0.41	0.505
rs4889616	78559	31155509	C/G			
rs1045564	79549	31156499	A/C			
rs2303221	79882	31156832	C/T			
rs1549296	81339	31158289	A/G			
rs889555	81681	31158631	C/T	0.49	0.50	0.740
rs5816521	81696	31158646	-G			
rs749767	83517	31160467	C/T	0.28	0.36	0.020
rs2884738	85431	31162381	A/C			
rs2052581	86332	31163282	C/T			
rs4889617	87358	31164308	A/G			
rs4889619	87725	31164675	C/T			

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1978487	89052	31166002	A/G	<b>0.62</b>	<b>0.57</b>	<b>0.124</b>
rs1978486	90020	31166970	A/G			
rs1978485	90231	31167181	A/G	<b>0.90</b>	<b>0.88</b>	<b>0.513</b>
rs4889620	90284	31167234	A/G			
rs4889621	90447	31167397	C/T			
rs3214477	90601	31167551	-G			
rs4527034	90724	31167674	A/G	<b>0.37</b>	<b>0.43</b>	<b>0.079</b>
rs1060506	92559	31169509	C/T	<b>0.29</b>	<b>0.28</b>	<b>0.720</b>
rs7200125	95176	31172126	A/G			
rs6565219	95195	31172145	C/T			
rs889548	96822	31173772	A/G	<b>0.54</b>	<b>0.51</b>	<b>0.320</b>
rs6145813	Not mapped	Not mapped	/TTTTTT TTTTTT	<b>0.33</b>	<b>0.32</b>	<b>0.909</b>

[0230] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1A for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1C can be determined by consulting Table 13. For example, the left-most X on the left graph is at position 31077197. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0231] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than  $10^{-8}$  were truncated at that value.

[0232] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken

horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

### Example 5

#### Chromosome 4 Region Proximal SNPs

[0233] SNP rs1957723 is associated with osteoarthritis and is described in Table A. One hundred-thirty additional allelic variants proximal to rs1957723 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 14. The chromosome positions provided in column four of Table 14 are based on Genome "Build 34" of NCBI's GenBank.

TABLE 14

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs3849023	4	211	36870611	g/t
rs1444311	4	7217	36877617	a/g
rs2044295	4	7895	36878295	a/c
rs2166093	4	13308	36883708	c/t
rs2376334	4	14279	36884679	g/t
rs1444320	4	17026	36887426	c/t
rs2044294	4	18271	36888671	a/g
rs1899864	4	20417	36890817	c/t
rs1562094	4	21843	36892243	a/g
rs1562098	4	22069	36892469	a/g
rs1562097	4	22145	36892545	a/g
rs1562096	4	22519	36892919	a/g
rs1562095	4	22539	36892939	a/g
rs1444319	4	23236	36893636	a/c
rs1444318	4	23256	36893656	a/c
rs1025938	4	23402	36893802	c/t
rs1025937	4	23499	36893899	a/c
rs1025936	4	23620	36894020	c/t
rs1020333	4	23871	36894271	a/t
rs2120654	4	24136	36894536	c/g
rs2588547	4	25427	36895827	a/g
rs2044293	4	25866	36896266	g/t
rs2760324	4	26541	36896941	a/g
rs2588546	4	26576	36896976	g/t
rs2588545	4	26689	36897089	a/g
rs2760328	4	26720	36897120	a/c
rs2588544	4	27113	36897513	c/t
rs2760331	4	27164	36897564	c/t
rs2588543	4	27186	36897586	a/g
rs2588542	4	28341	36898741	a/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs2588541	4	29160	36899560	c/t
rs2588540	4	29844	36900244	a/g
rs2760336	4	30665	36901065	g/t
rs2760337	4	30830	36901230	a/g
rs2028732	4	31061	36901461	a/c
rs2588538	4	31523	36901923	c/t
rs1992617	4	32326	36902726	c/t
rs1998469	4	32346	36902746	a/g
rs1998470	4	32358	36902758	c/t
rs1975498	4	34909	36905309	c/t
rs1562093	4	34975	36905375	a/g
rs1975497	4	35066	36905466	c/t
rs1562092	4	35096	36905496	g/t
rs2248788	4	35375	36905775	c/t
rs1899862	4	36304	36906704	a/g
rs2588532	4	36712	36907112	a/t
rs1885878	4	36770	36907170	c/t
rs986648	4	37342	36907742	c/t
rs986647	4	37412	36907812	c/t
rs1010010	4	37884	36908284	a/g
rs1010009	4	38077	36908477	a/c
rs2760325	4	38300	36908700	c/t
rs2588531	4	38301	36908701	c/t
rs1838388	4	41189	36911589	c/t
rs1975495	4	44408	36914808	c/t
rs2181491	4	44493	36914893	a/c
rs1975496	4	44571	36914971	a/g
rs2181492	4	44670	36915070	a/g
rs2224719	4	45219	36915619	a/g
rs2224720	4	45258	36915658	c/t
rs1951770	4	47261	36917661	a/g
rs2296040	4	48473	36918873	a/c
rs1957723	4	48771	36919171	a/g
rs1957725	4	55292	36925692	c/t
rs2889346	4	56479	36926879	a/g
rs1885879	4	56747	36927147	a/c
rs1957726	4	60620	36931020	g/t
rs1957727	4	60688	36931088	a/c
rs1885880	4	61058	36931458	a/c
rs1885881	4	61129	36931529	c/t
rs942108	4	61577	36931977	c/t
rs1951771	4	61961	36932361	a/g
rs2376323	4	63351	36933751	g/t
rs2013358	4	63926	36934326	a/g
rs2181494	4	65798	36936198	a/g
rs1957728	4	66043	36936443	a/c
rs1957729	4	66044	36936444	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs1957730	4	66246	36936646	c/t
rs1957731	4	66318	36936718	c/t
rs1998468	4	66547	36936947	g/t
rs1957732	4	71238	36941638	c/t
rs1957733	4	71283	36941683	a/g
rs2376322	4	71492	36941892	a/g
rs2889345	4	72274	36942674	a/g
rs1815267	4	73762	36944162	a/t
rs1957734	4	74209	36944609	g/t
rs1957735	4	75284	36945684	a/t
rs1957736	4	77347	36947747	a/c
rs1957737	4	77589	36947989	c/t
rs1957738	4	78096	36948496	a/g
rs1957739	4	78606	36949006	a/g
rs1957740	4	78862	36949262	g/t
rs1957741	4	79135	36949535	a/g
rs1957742	4	79146	36949546	a/g
rs1957743	4	79456	36949856	c/t
rs1957744	4	79609	36950009	a/g
rs1957745	4	80086	36950486	a/g
rs1957746	4	80119	36950519	a/g
rs1957747	4	80766	36951166	c/t
rs2146670	4	81110	36951510	a/g
rs2146671	4	81269	36951669	a/t
rs1957748	4	81668	36952068	c/t
rs2162307	4	82433	36952833	c/t
rs1962839	4	82559	36952959	c/g
rs2376315	4	83298	36953698	c/t
rs1426410	4	83821	36954221	a/g
rs1895921	4	84121	36954521	c/t
rs1895922	4	84147	36954547	c/t
rs1035779	4	84543	36954943	a/g
rs1035780	4	84554	36954954	a/g
rs1035781	4	84691	36955091	a/g
rs1035782	4	84727	36955127	a/g
rs1426411	4	85678	36956078	c/t
rs1834602	4	86699	36957099	c/t
rs1834603	4	86700	36957100	a/g
rs1834604	4	86792	36957192	a/g
rs1834605	4	86832	36957232	a/g
rs2162308	4	87045	36957445	a/g
rs1365341	4	87140	36957540	a/g
rs1820458	4	87365	36957765	a/c
rs1469310	4	88342	36958742	c/t
rs3057879	4	88498	36958898	-t/ca
rs1469311	4	88589	36958989	a/g
rs768326	4	95502	36965902	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs1863523	4	96968	36967368	c/t
rs1469312	4	97448	36967848	c/t
rs1469313	4	97568	36967968	c/t
rs1951773	4	98724	36969124	c/t
rs2120655	4	Not mapped	Not mapped	t/g
rs2181495	4	Not mapped	Not mapped	g/a

Assay for Verifying and Allelotyping SNPs

[0234] The methods used to verify and allelotype the 130 proximal SNPs of Table 14 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 15 and Table 16, respectively.

TABLE 15

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs3849023	ACGTTGGATGGGTAATTGCTAACCATGTTT	ACGTTGGATGGACCCAGTCAGTCAATAAAC
rs1444311	ACGTTGGATGGCCTATTGGTTTAACTAGGC	ACGTTGGATGTCCTGGCTTCTCAGGAGTTC
rs2044295	ACGTTGGATGCCACACCACACTACTATTCAAG	ACGTTGGATGGTGGTGTGTTAGAGAGGTTAC
rs2166093	ACGTTGGATGAAAACTCGTGAGATGGATGG	ACGTTGGATGTAGGTGTACAGTTCAGTGTC
rs2376334	ACGTTGGATGCTCTCAGAGAACCAGCTTTTG	ACGTTGGATGGGGAATATTAACAATTGGGG
rs1444320	ACGTTGGATGTAATTCTCTCCTCCAAATGC	ACGTTGGATGCTAGAAACAAAAGACTACATG
rs2044294	ACGTTGGATGAACCTAAATCTCCTCAAGCC	ACGTTGGATGTTCTGACCACTTCTCTATGG
rs1899864	ACGTTGGATGTTTATAGGCGTGGGCAATCG	ACGTTGGATGTTGTCAGAAAGTGTCTGTGCC
rs1562094	ACGTTGGATGTGGATTCCTTTCTGAAGAC	ACGTTGGATGGCAACAAAGAACTTAATGC
rs1562098	ACGTTGGATGCTCTGAGTCCGAGTGATCATC	ACGTTGGATGAAACAATTAGCAGGGCACAG
rs1562097	ACGTTGGATGCACAGGATCTTACTCTGTTG	ACGTTGGATGCGGACTCAGAAATTCAGTC
rs1562096	ACGTTGGATGACCCAGGGCATGTTATATAG	ACGTTGGATGTTTCTCTCTGGTACCCTCTC
rs1562095	ACGTTGGATGTGTTAGTAAACCAGGGCATG	ACGTTGGATGTGACAGATGCCACCAGTTAC
rs1444319	ACGTTGGATGTTCAACTTTAGCCTCTGGGC	ACGTTGGATGGGTCAGAAAGTCAAGGAAC
rs1444318	ACGTTGGATGCTCTGGGCAATATCAAGCC	ACGTTGGATGATGCTCGCTGATGTTGTTGGG
rs1025938	ACGTTGGATGCAGGTAAAGAAAGCTTTTGG	ACGTTGGATGCCCTCAGTAACTAGTGAATTC
rs1025937	ACGTTGGATGGAATAGGAAAGGTAGTATACC	ACGTTGGATGAAATTCAGTCATTAGCAGGGG
rs1025936	ACGTTGGATGTCTCCAGGTAGATGAGTCAG	ACGTTGGATGCCACACACCAAGCAATCAC
rs1020333	ACGTTGGATGGCATCTCTCAATCTGGACG	ACGTTGGATGGTGGATCAGCAAGTCAGAG
rs2120654	ACGTTGGATGACCAGAAAGACCAGGGCATG	ACGTTGGATGAACCTTTAGCTCTTCTCCCC
rs2588547	ACGTTGGATGTACAAATGTAATATAATC	ACGTTGGATGGATAGCTACGTTTAAAAATG
rs2044293	ACGTTGGATGTGTCAACAATACAAAGACTAA	ACGTTGGATGTGCACTGGACCTTTTTTTT
rs2760324	ACGTTGGATGACAAACCAGTGGTTGAGGAG	ACGTTGGATGCCTCAGCAATCCAACAGAAC
rs2588546	ACGTTGGATGCTTAGAGGATGGAGTCAGTC	ACGTTGGATGTACTACCAGAGATGCTGGTG
rs2588545	ACGTTGGATGCAACACAGCTACAGTGCATC	ACGTTGGATGTGGGTAAAGGGGAAAGGAAGG
rs2760328	ACGTTGGATGGCCATAAAATTTGGGTAAGGG	ACGTTGGATGGCATCTATTTGACACCAACG
rs2588544	ACGTTGGATGTAAGAATTAGCATGTGAAAG	ACGTTGGATGTTTGTGCACAAAGAAATTTGG
rs2760331	ACGTTGGATGAACAGTATGCCCTTTTGTGC	ACGTTGGATGCTCTCGTAATTTTACATGAC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2588543	ACGTTGGATGGTGGCAAAATCTTTGTGCAC	ACGTTGGATGCTAAAGTAGGTAGTATACCAG
rs2588542	ACGTTGGATGTGGCAGCAAGGCTTAAGCTC	ACGTTGGATGTCACAGTCACCTCTCATTC
rs2588541	ACGTTGGATGTGAACAAGGCTATGTCAGGG	ACGTTGGATGGGCATTGTCTAGGTGATGAG
rs2588540	ACGTTGGATGTGCTGTATGATCCAGCAATC	ACGTTGGATGGTGCAAAATCTGCTCTCTC
rs2760336	ACGTTGGATGAAGCTGAGGCAGGAGAATGG	ACGTTGGATGTGTTTTGAGACGGAGTCTCG
rs2760337	ACGTTGGATGGGGTGTTCGAACTGATACAAG	ACGTTGGATGACTACCATTCTACTCTCTCG
rs2028732	ACGTTGGATGTTCTCGGACAGCTAAATAGG	ACGTTGGATGGCCATTGCTGTTTTCTGTGT
rs2588538	ACGTTGGATGTATCTTCTGGGAAGCCTTTC	ACGTTGGATGGACTTGAATCACTCCATGC
rs1992617	ACGTTGGATGGGAGGACATTGCCCTCAAAG	ACGTTGGATGCTGACCTTCTGTCTAGTCAC
rs1998469	ACGTTGGATGTATATGCCAAGGACCAACGG	ACGTTGGATGCTGACCTCTGTCTAGTCAC
rs1998470	ACGTTGGATGATTCCCCCAATTAAAGCTTTG	ACGTTGGATGGAAAAGTATTATATGCCAAGG
rs1975498	ACGTTGGATGAGCTCTCTTTTGCCTGCTG	ACGTTGGATGAGGAGGCTTCAACATCATGG
rs1562093	ACGTTGGATGTGATGTGTGAAGCCTCCTCTG	ACGTTGGATGAAAGACATACCCCAAGACTGG
rs1975497	ACGTTGGATGTCAAGCAGCATGAAAACGAC	ACGTTGGATGCACTTTAGACTTTTCTGGGG
rs1562092	ACGTTGGATGTTCCAGTGACTGGACCATAG	ACGTTGGATGTCAGCAGCATGAAAACGTAC
rs2248788	ACGTTGGATGGGGGAAAGGAAAAAGACTTCC	ACGTTGGATGTGCTAGTAGTGGCTCTCAAAG
rs1899862	ACGTTGGATGTAATCTCCCAATAAAGTAGG	ACGTTGGATGGCTACAAAAGAAAATGAATAC
rs2588532	ACGTTGGATGCAACAAATAGTGGCTGAGAG	ACGTTGGATGTTTGTACACAGCCGCTATAG
rs1885878	ACGTTGGATGTGACTCAGCGAGTTTGATAGC	ACGTTGGATGAGCCAGATTGGGTGCTTTTC
rs986648	ACGTTGGATGTGAGAAAGCTTTCTGAGGAC	ACGTTGGATGGGTTTTCTGTTGTGAATGGG
rs986647	ACGTTGGATGACACACTCTTTCTCAAGCAG	ACGTTGGATGCTTATTTGTCTCAGAAAGC
rs1010010	ACGTTGGATGACTGTAGCTAAGTTGGCAT	ACGTTGGATGTTCAACCAACCAATAAGGC
rs1010009	ACGTTGGATGTCTCATCAGCTCTTTCTCTGG	ACGTTGGATGAAAGGGATGAGGAAGTGAGG
rs2760325	ACGTTGGATGATCCCCAGCATGTAGCATAG	ACGTTGGATGCTGCCCAATAGTCTCTCTG
rs2588531	ACGTTGGATGATCCCCAGCATGTAGCATAG	ACGTTGGATGCTGCCCAATAGTCTCTCTG
rs1838388	ACGTTGGATGGTACCTCATGGATATTTACAC	ACGTTGGATGTGGTGTTGTATAAATGAC
rs1975495	ACGTTGGATGCAAGTGACGAGTTTAAAGACC	ACGTTGGATGAGCTGGGATTACAGTCATGC
rs2181491	ACGTTGGATGGTACCTAATATATGCTTCTGG	ACGTTGGATGTTATTCCTGCTTACTTTTCC
rs1975496	ACGTTGGATGTATATTAGTACAGTTGTGGC	ACGTTGGATGCAACCAACCTTCACTGAAAGC
rs2181492	ACGTTGGATGCTTTGAGGAAGAGGAAGAAG	ACGTTGGATGCAACTACCTTTTGGAGGCAG
rs2224719	ACGTTGGATGTCAAGGGGTAGATGTGTAG	ACGTTGGATGCCAGAGAGGAGTAATGGTAT
rs2224720	ACGTTGGATGCCAATTACTCAAGGGGTGAG	ACGTTGGATGAAATTCAGTACAGCAGAGGG
rs1951770	ACGTTGGATGCCTGGGAACCTCAGCTTTTC	ACGTTGGATGTGGCACAGCAGGAATATCAG
rs2296040	ACGTTGGATGGGGCATTATGAAATGCAGAC	ACGTTGGATGGCATTGTACAGGAAGAAGCATG
rs1957723	ACGTTGGATGTACTCACCTTGTGACTGCTC	ACGTTGGATGGCTGACGGTACACATTAATC
rs1957725	ACGTTGGATGTTATTGGAATCTCCAGGTC	ACGTTGGATGAAGATGATTAGTCCAGCCTG
rs2889346	ACGTTGGATGTGACTGACTTCCTAGGTCAG	ACGTTGGATGTGACAGTGTGTTGAGTGGCAG
rs1885879	ACGTTGGATGTCTCACCCCTTACATCTGAT	ACGTTGGATGCTACAGGAAGATGACAGAG
rs1957726	ACGTTGGATGAAATTCAGCCACTCAACCAG	ACGTTGGATGAAGTGGTTGGGATTGTGAG
rs1957727	ACGTTGGATGGCCAACGATATCTTAAACCC	ACGTTGGATGGTTTGTCTGGTTCTCATC
rs1885880	ACGTTGGATGTGAGTGGCCCAAGATTTCAC	ACGTTGGATGCTGGAATCCCAAGGTTCTCTG
rs1885881	ACGTTGGATGTAGACGTGTTCTGCATCATG	ACGTTGGATGATGAAATCTTGGGGCATTCC
rs942108	ACGTTGGATGGAGCTGTTAGGGTAGAAATG	ACGTTGGATGGTCTTGGACTAATTTTGACC
rs1951771	ACGTTGGATGGGCATTCTCCTTTTGTCTAAG	ACGTTGGATGAGTAAACCAAGGACTAGAGCC
rs2376323	ACGTTGGATGTGCTTACTTGTCTAGCACTGC	ACGTTGGATGGCATCCCTTGGTGATGATA
rs2013358	ACGTTGGATGGGAAATTTAGGAGTACTGTAG	ACGTTGGATGGCCCAACCATGAAACCTAAATC
rs2181494	ACGTTGGATGATTCAATTACCTCCCACTGG	ACGTTGGATGTATCCCAACCCAAATGTCAC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs1957728	ACGTTGGATGAAATAGATCCCAACCAAGGG	ACGTTGGATGGTAAACATTACCTAAGCGGG
rs1957729	ACGTTGGATGAAATAGATCCCAACCAAGGG	ACGTTGGATGGTAAACATTACCTAAGCGGG
rs1957730	ACGTTGGATGGGTCTAAACATGAGAGACTC	ACGTTGGATGTCCTTTATGGATATAGGGTCC
rs1957731	ACGTTGGATGTATTGGAACTGGTACCTGG	ACGTTGGATGGACCTGAATCATGTCTCCAG
rs1998468	ACGTTGGATGTATAAGGCTCTAAAAGTGGG	ACGTTGGATGACCTTATCCAGAATGAAAC
rs1957732	ACGTTGGATGAAGAGAGGAGTTTATTGGCC	ACGTTGGATGCGGCGCTGATCTTTATTTTGG
rs1957733	ACGTTGGATGCTATCAAGACTCTGATTGCC	ACGTTGGATGTGTTTGCAGGTAACCTTGGC
rs2376322	ACGTTGGATGTCGTTCTCTCTGTGCATG	ACGTTGGATGTTAGTCAGATGCTTGTGTAG
rs2889345	ACGTTGGATGTGGAATCCCAACCTTTTACG	ACGTTGGATGTTCTTGCTAAATGTAGGGCC
rs1815267	ACGTTGGATGCGAGGAAGGGCTACTATCAG	ACGTTGGATGGTAGGCCAAACCTAGCTTTGG
rs1957734	ACGTTGGATGCTACCCCTGCCCTTATAATT	ACGTTGGATGCAAGTGGTAAAGGATGTGG
rs1957735	ACGTTGGATGAGCTTCCCATGGTTATAGAG	ACGTTGGATGCTGAAAACAATACCGGTCTC
rs1957736	ACGTTGGATGCTGAAGCAAAAGATTCTCTC	ACGTTGGATGAGCATCTTTTGGCTGACCTG
rs1957737	ACGTTGGATGACATGGAAGCTGAAAGCCAA	ACGTTGGATGCAAGAGCTTTGACCTTACCTC
rs1957738	ACGTTGGATGATGTCCCTTAAAGGCTGCC	ACGTTGGATGCAGATGATCTTGTCTCCAG
rs1957739	ACGTTGGATGCTACTGCCCTGAGTGCTTTAG	ACGTTGGATGCTGATGGCCTGAGAACTAAG
rs1957740	ACGTTGGATGGCCCAAGTCAAGTTGACATAC	ACGTTGGATGCACTGCTCCAGTTATATAC
rs1957741	ACGTTGGATGCTAGTGATACCCCTATTAG	ACGTTGGATGCCCTTAGTAAATATGGATG
rs1957742	ACGTTGGATGGGATGATATCTACTTTGTACG	ACGTTGGATGGACTCCATCTGAGATGTTAG
rs1957743	ACGTTGGATGCAACTGTCTTGTATTGAAAG	ACGTTGGATGGACAGACTTTGATTTTTC
rs1957744	ACGTTGGATGTCAAGTGATACCCGTGTAATGCC	ACGTTGGATGTGCCCAAGCAGTGAATATTG
rs1957745	ACGTTGGATGGTTTGAAGAAAGTGTGGGTTCC	ACGTTGGATGAGCAATGACGCTTATTACC
rs1957746	ACGTTGGATGCTCTCATACACATAGTTAGC	ACGTTGGATGGGTTTAGGTTTGGTTTGGT
rs1957747	ACGTTGGATGGTCACTCAAGATAACAGTTCC	ACGTTGGATGTACCTAACCTGGTGAAGGTAGC
rs2146670	ACGTTGGATGCCTAACACATCTTTATGAGC	ACGTTGGATGCTCATAGATATGCTAAGCAC
rs2146671	ACGTTGGATGATGAGGAGCACTAGAAGGC	ACGTTGGATGAAAGGGCTGGAAGAAACAGG
rs1957748	ACGTTGGATGTGAAGTTTGTAGTAGGGAGC	ACGTTGGATGTTCTGTGCACACAACACTCC
rs2162307	ACGTTGGATGACATGCGGTGCTGGCCCTTT	ACGTTGGATGCTTTGTAGAGGACATGGATG
rs1962839	ACGTTGGATGGGCTGCATAGTATTCATGG	ACGTTGGATGAGGGAATCCCTTCCCCATTG
rs2376315	ACGTTGGATGTGGCCTTGGATTCTTCCAC	ACGTTGGATGAGAATTGGACAGAGTGGCAG
rs1426410	ACGTTGGATGGAGAAAGTTGCATCTTGCC	ACGTTGGATGGGGAAGTTTACCTTGGCTC
rs1895921	ACGTTGGATGGGTGATGGTGTGGAGGTAC	ACGTTGGATGATTAGGCTTGTCCCAACATC
rs1895922	ACGTTGGATGCAATGCATTAGGCTCTCTCC	ACGTTGGATGGAGGTACATTTCTCAGGCAG
rs1035779	ACGTTGGATGGAGAATCACTTGAACCCGGG	ACGTTGGATGTGGAAGTGAGTGGGATGATC
rs1035780	ACGTTGGATGTTTGAAGTGGAGTCTCGCTC	ACGTTGGATGATCAACTTGAACCCGGGAGG
rs1035781	ACGTTGGATGGGAAGATGCTGACTCTGAAC	ACGTTGGATGCCCTGACTGTTTAGGGATCC
rs1035782	ACGTTGGATGGGATCCCTAAACAGTCAAGG	ACGTTGGATGAGTTGGCTAGACTTGCCTTC
rs1426411	ACGTTGGATGGAGAGTGCATCAACAGTCG	ACGTTGGATGTGATCCTTGGTCAAGGTGATC
rs1834602	ACGTTGGATGGATGGGCCCTATTTTCTTG	ACGTTGGATGCTTTTCCAACCCAGTAATGTC
rs1834603	ACGTTGGATGGATGGGCCCTATTTTCTTG	ACGTTGGATGCTTTTCCAACCCAGTAATG
rs1834604	ACGTTGGATGGAGAAAGACATTACTGGGTTGG	ACGTTGGATGAGAATCTTCCGTGAGCTGGG
rs1834605	ACGTTGGATGGCCCAACAGTCAGGAAGAAAT	ACGTTGGATGTTTGGAGAGCTGGCCAAAAG
rs2162308	ACGTTGGATGTAAACAAACAGAGGACACC	ACGTTGGATGTATGATCAGAGTCAAGCGG
rs1365341	ACGTTGGATGTCCCTCTGTTCTTTAGGCA	ACGTTGGATGCATCTCCCTGGTATGCAATT
rs1820458	ACGTTGGATGCACCTCAGACTTGGAAATG	ACGTTGGATGGTCAAGTGATGCTATTACAGC
rs1489310	ACGTTGGATGTACTACAGCGTGTTTAGCAG	ACGTTGGATGTGTCAAAGGAGGATGTAGAG
rs3057879	ACGTTGGATGGGCACATTGGAATAAAGCC	ACGTTGGATGACGGCATGAACAATTCTCAG



dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs1469311	ACGTTGGATGCCTGAGAAATTGTCATGCCG	ACGTTGGATGTTTTCAGTGTCTCTCCAGG
rs768326	ACGTTGGATGAATTAGCCAGGCATGGTGTC	ACGTTGGATGACATCCTAGGCTCAAGTGAC
rs1863523	ACGTTGGATGGGCAGACACATTCCTATTTCG	ACGTTGGATGGGGAAGGTGTGCTGAGTAA
rs1469312	ACGTTGGATGCATTCTGTCAGCAATCTAGC	ACGTTGGATGGGACTCATGTCTCTCTGG
rs1469313	ACGTTGGATGAGTGAGGGAGAAAAGTGAAC	ACGTTGGATGCCTAACTTCTCTCCAATCTC
rs1951773	ACGTTGGATGAAGGTTCAAGTTACCCGATG	ACGTTGGATGCACTGTGGTCCATGAAAAA
rs2120655	ACGTTGGATGACAGGGTTTCTGCATGTTGC	ACGTTGGATGACGCCTGTAATCCAGCACT
rs2181495	ACGTTGGATGGAATTGTGGGAGTTACAATTC	ACGTTGGATGGAATCAAGCTAATTAACATGTG

TABLE 16

dbSNP rs#	Extend Primer	Term Mix
rs3849023	CTCATAACATAAGAAGTTGATGC	CGT
rs1444311	CTAGGCATGCTAGCTTGCC	ACT
rs2044295	CACTACTATTCAAGATTACCCITT	ACT
rs2166093	GGTGGTGATGGCTGCACAA	ACG
rs2376334	TCAGAGAACCAGCTTTTGATTTC	ACT
rs1444320	GCCTAGACCCCGTGCAAC	ACG
rs2044294	CTCCTCAAGCCAATAGGTCTTA	ACG
rs1899864	CGCACCTGGCCGAAAATAAC	ACT
rs1562094	AACCTGCAAAAAGATTACACTTGC	ACT
rs1562098	TCCTGCCTCAGCCTTCTCTAGA	ACT
rs1562097	ACTCTGTTGTTTCAGGCTGGGT	ACT
rs1562096	TAAGCTAGCTAGTAAGTGGTG	ACT
rs1562095	ATGTTATATAGAACATCCCTTTTT	ACT
rs1444319	TCTGGGCAATTATCAAGCCTTT	ACT
rs1444318	CTTTGCATTTTCTGAGTTTCTTT	ACT
rs1025938	AAGAAAAGCTTTTTGGTTTGGG	ACT
rs1025937	GGTAGTATACCTAAAAAACAGC	CGT
rs1025936	TCAAAGGACACCCAGCATTCA	ACG
rs1020333	ACGTTTATCTGTAACTTTTCCA	CGT
rs2120654	GAAAGACCAGGGCATGATTAGA	ACT
rs2588547	ACAAATGTAATATAAATCAAGCTC	ACG
rs2044293	ACCAGCCTGGGTAAACATAGCCA	ACT
rs2760324	GGTTGAGGAGAAGCACCAGCA	ACG
rs2588546	TACAATTTCTAGCCTTAATAAGAT	ACT
rs2588545	TACAGTGCATCTATTTGACACCAA	ACG
rs2760328	AAATTGGGTAAGGGAAAAAG	ACT
rs2588544	ATTAGCATGTGAAAGACTTCTC	ACT
rs2760331	AGTATGCCTTTTGTGCACAAAGA	ACT
rs2588543	ATTCTTTGTGCACAAAAGGCATA	ACG
rs2588542	GCTTAAGCTCTTACAGGCAG	CGT
rs2588541	AGGTCTATGTCAGGGAAAACCTTA	ACG

dbSNP rs#	Extend Primer	Term Mix
rs2588540	GATCCAGCAATCCCAGTGAT	ACG
rs2760336	AGGCGGAGCTTGCAAGTGAG	ACT
rs2760337	CACCAATACGTGATGATCTTTT	ACT
rs2028732	CAGCTAAATAGGGCTTGAGTCAAT	CGT
rs2588538	AATTTGTACAAATTTATGGGGTAT	ACT
rs1992617	ATTGCCTTCAAAGAACATCAAAGC	ACG
rs1998469	GACCAACGGGAGGACATTG	ACG
rs1998470	CTTTGAAGGCAATGTCCTCC	ACG
rs1975498	TTTTGCCTGCTGCTATCCAC	ACT
rs1562093	CTCCTCTGCCATGTGGAAAC	ACG
rs1975497	AAAACGACTAATACACACTGTT	ACT
rs1562092	TTTGGTTAATGGACATTAGACT	ACT
rs2248788	TGTGGGATTTTATTATTTTCATCA	ACT
rs1899862	TAAGTGCATAACTTGTCTTGAGG	ACT
rs2588532	ATAGTGGCTGAGAGCCAGAT	CGT
rs1885878	GCGAGTTTGTAGCACAGGC	ACT
rs986648	GTACATGTAATGCTAGTAAAGAAA	ACG
rs986647	CTCTTTCTCAAGCAGGAGTTA	ACG
rs1010010	AGCTAAGTTGGCATGTGGGA	ACT
rs1010009	CCTGGCTACCTTCCAAAAAG	ACT
rs2760325	TCTCAGGAAGATGAAAATAATAG	ACG
rs2588531	CTCAGGAAGTATGAAATAAATAGT	ACT
rs1838388	TCATGGATATTACACCTACTAC	ACT
rs1975495	AGGAGTTTAAGACCAGCCTG	ACT
rs2181491	TGCTTCTGGATTTTAAATGATCAC	ACT
rs1975496	ATGATCAATCATTTTGAGGGC	ACT
rs2181492	GTTGCATTGCTATGGTCTGC	ACT
rs2224719	CATATATCCCTCTGTCTGTAC	ACG
rs2224720	GGGTGTAGATGTGTAGATTATA	ACT
rs1951770	ACAAGCATTAGAGACTTGATTG	ACG
rs2296040	CTTTGTTTCTAAATCTGATAGTC	ACT
rs1957723	AGCATGGCATAGGCACTGG	ACG
rs1957725	GCGAGGAAAAGACCTGTCTA	ACG
rs2889346	GGTCAGCTCAGCTGGTTTTT	ACG
rs1885879	CACATCTGATGCTCTCCTAAA	ACT
rs1957726	GCCACTCAACCAGTAGGAAA	ACT
rs1957727	GTATCTTTAAACCCCTCACAAAT	CGT
rs1885880	TTACGTTAGTCTGCTACTTCCA	ACT
rs1885881	TGGGCTATCAATGATGGAAC	ACT
rs942108	AAATGAAATAGAATTGTGACTTC	ACT
rs1951771	TCCTTTTGTCTAAGAATATTAG	ACG
rs2376323	CTAGCACTGCCAAGTGCAAC	ACT
rs2013358	TTTTAGGAGTACTAGAACACA	ACG
rs2181494	TGGTCCCTGCCATAACAC	ACT

dbSNP rs#	Extend Primer	Term Mix
rs1957728	AGAAGCATGTGCTTATAACAATAA	CGT
rs1957729	GAAGCATGTGCTTATAACAATAAAA	ACT
rs1957730	ACATGAGAGACTCTGAAGACT	ACT
rs1957731	GGGTGAGCTTTGGGATCAC	ACT
rs1998468	GGGCATAATTAATCCATGTTAG	ACT
rs1957732	GGCCAAGTTACCTGCAAAC	ACT
rs1957733	TCTAATGTTAAAGAGAGGAGTTTA	ACG
rs2376322	GCGCCAAGGAAAGGCCAC	ACT
rs2889345	TCATTTCTCACCCCTTGATATCCA	ACT
rs1815267	AAAGGGCTACTATCAGTTTTGT	CGT
rs1957734	CTGCCCTATAATTCTAAAAAGGT	ACT
rs1957735	CTAAACTAAGAAATGTTCCAC	CGT
rs1957736	TAATACTAAGGAGAGGGCTCCT	ACT
rs1957737	AGCCAAGGGTGTGGATGAG	ACT
rs1957738	CCTTAAAGGGCTGCCTACAAAATA	ACT
rs1957739	CTGAGTGCTTTAGCTGGATTA	ACG
rs1957740	TTAAGCATCACACTGAGTTTGAG	ACT
rs1957741	AGCTGAATTAAGCGGACAGCTA	ACG
rs1957742	TCTACTTTGTACGTAGCTGTCCG	ACT
rs1957743	GAAAAATTACTAAAAAGACCTC	ACG
rs1957744	TGTACCTGTAAATGCCTAAAGC	ACG
rs1957745	TTTTCAAAGGTTTAGGTTTGGTTT	ACT
rs1957746	ACAACATAGTTAGCAATGCAG	ACG
rs1957747	GATAACAGTTCCAATTACAACAA	ACG
rs2146670	ATCTTTATGAGCTTTTCCCTTCTT	ACG
rs2146671	TACAACCCCTTTCAGGACTTCA	CGT
rs1957748	TTGTAGTAGGGAGCCATGGT	ACT
rs2162307	CCTGGCCCTTTGTCCCTG	ACG
rs1962839	CCACATCTTTGACAAACCTGA	ACT
rs2376315	CCCCTTCTTTTCCAGGC	ACT
rs1426410	CATCTTGCCCTAAATCACTC	ACG
rs1895921	GTACATTTCTCAGGCAGCTC	ACG
rs1895922	ATTAGGCTTCTCCACCATC	ACT
rs1035779	ACCCGGGAGGGTTGCAGT	ACT
rs1035780	GGCTGGAGTGCACTGGCA	ACG
rs1035781	ACCTAGACTAAGAGAGTGATTGCA	ACT
rs1035782	CCTAAACAGTCAAGGCAAAGG	ACT
rs1426411	TTTATGGCTTCTTAGGATATCA	ACG
rs1834602	AGGAAGGTGCCAGATCCT	ACG
rs1834603	AGGAAGGTGCCAGATCCTT	ACT
rs1834604	AGTTTTCTAGTAACCTTCTCTAAA	ACT
rs1834605	ACAGTCAGGAAGAATTCTGTCT	ACT
rs2162308	CACCTACAGAGTTTAAAGTAAATTT	ACG
rs1365341	AAATCTCCTGGAGGGCTTCATAA	ACT

dbSNP rs#	Extend Primer	Term Mix
rs1820458	TGGAAATGGCAACTGAATCCT	ACT
rs1469310	ACCCACACAAATGCCAATAGCAC	ACT
rs3057879	TGGAAATAAAGCCTTTTGAGGTT	ACT
rs1469311	TGCCGTTAAAGAGGAAAAGCT	ACT
rs768326	CAGCTACTCTGTAAAGCTGAA	ACT
rs1863523	ATATCTTGCTCATCTTTCTCTAT	ACT
rs1469312	TAGTCCAGCAAAAGCCAGC	ACT
rs1469313	GTGAACAAATAATGCAAGTTCAG	ACT
rs1951773	CCCTTTGGGAGAGAAGGGC	ACT
rs2120655	AGCAATCCTCCCACTTTGGC	CGT
rs2181495	GGTGACATTGGGTGGGATACA	ACT

### Genetic Analysis

[0235] Allelotyping results from the discovery cohort are shown for cases and controls in Table 17. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where "AF" is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs1444311 has the following case and control allele frequencies: case A1 (A) = 0.74; case A2 (G) = 0.26; control A1 (A) = 0.75; and control A2 (G) = 0.25, where the nucleotide is provided in parenthesis. Some SNPs are labeled "untyped" because of failed assays.

TABLE 17

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs3849023	211	36870611	G/T			
rs1444311	7217	36877617	A/G	0.26	0.25	0.566
rs2044295	7895	36878295	A/C			
rs2166093	13308	36883708	C/T			
rs2376334	14279	36884679	G/T	0.15	0.16	0.734
rs1444320	17026	36887426	C/T			
rs2044294	18271	36888671	A/G	0.16	0.14	0.412
rs1899864	20417	36890817	C/T			
rs1562094	21843	36892243	A/G	0.22	0.23	0.586
rs1562098	22069	36892469	A/G			
rs1562097	22145	36892545	A/G	NA	0.97	NA
rs1562096	22519	36892919	A/G	0.20	0.21	0.773
rs1562095	22539	36892939	A/G	0.53	0.51	0.407
rs1444319	23236	36893636	A/C	0.74	0.79	0.023
rs1444318	23256	36893656	A/C	0.12	0.13	0.559
rs1025938	23402	36893802	C/T	0.18	0.19	0.633
rs1025937	23499	36893899	A/C			
rs1025936	23620	36894020	C/T	0.84	0.84	0.907
rs1020333	23871	36894271	A/T			
rs2120654	24136	36894536	C/G	0.15	0.16	0.718
rs2588547	25427	36895827	A/G	0.39	0.40	0.603
rs2044293	25866	36896266	G/T			

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2760324	26541	36896941	A/G	0.59	0.61	0.395
rs2588546	26576	36896976	G/T	0.07	0.05	0.352
rs2588545	26689	36897089	A/G			
rs2760328	26720	36897120	A/C	0.25	0.26	0.791
rs2588544	27113	36897513	C/T			
rs2760331	27164	36897564	C/T	0.91	0.94	0.184
rs2588543	27186	36897586	A/G	0.59	0.59	0.828
rs2588542	28341	36898741	A/T			
rs2588541	29160	36899560	C/T	0.61	0.59	0.313
rs2588540	29844	36900244	A/G	0.62	0.62	0.999
rs2760336	30665	36901065	G/T			
rs2760337	30830	36901230	A/G	0.16	0.16	0.826
rs2028732	31061	36901461	A/C	0.60	0.58	0.432
rs2588538	31523	36901923	C/T	0.62	0.61	0.853
rs1992617	32326	36902726	C/T	0.61	0.59	0.282
rs1998469	32346	36902746	A/G			
rs1998470	32358	36902758	C/T	0.81	0.86	0.018
rs1975498	34909	36905309	C/T			
rs1562093	34975	36905375	A/G	0.89	0.87	0.529
rs1975497	35066	36905466	C/T	0.13	0.13	0.691
rs1562092	35096	36905496	G/T			
rs2248788	35375	36905775	C/T	0.29	0.31	0.368
rs1899862	36304	36906704	A/G	0.18	0.16	0.274
rs2588532	36712	36907112	A/T	0.30	0.32	0.443
rs1885878	36770	36907170	C/T	0.35	0.35	0.866
rs986648	37342	36907742	C/T	0.74	0.73	0.679
rs986647	37412	36907812	C/T	0.78	0.76	0.263
rs1010010	37884	36908284	A/G	0.25	0.26	0.649
rs1010009	38077	36908477	A/C	0.26	0.25	0.781
rs2760325	38300	36908700	C/T			
rs2588531	38301	36908701	C/T			
rs1838388	41189	36911589	C/T	0.75	0.74	0.650
rs1975495	44408	36914808	C/T			
rs2181491	44493	36914893	A/C	0.14	0.12	0.235
rs1975496	44571	36914971	A/G	0.26	0.26	0.944
rs2181492	44670	36915070	A/G	0.11	0.09	0.311
rs2224719	45219	36915619	A/G	0.78	0.78	0.866
rs2224720	45258	36915658	C/T	0.20	0.21	0.641
rs1951770	47261	36917661	A/G	0.22	0.18	0.029
rs2296040	48473	36918873	A/C	0.41	0.43	0.459
rs1957723	48771	36919171	A/G	0.42	0.38	0.113
rs1957725	55292	36925692	C/T	0.75	0.78	0.196
rs2889346	56479	36926879	A/G	0.54	0.55	0.677
rs1885879	56747	36927147	A/C	0.44	0.48	0.123
rs1957726	60620	36931020	G/T	0.14	0.14	0.741
rs1957727	60688	36931088	A/C	0.73	0.76	0.271
rs1885880	61058	36931458	A/C	0.43	0.43	0.935
rs1885881	61129	36931529	C/T	0.12	0.11	0.681
rs942108	61577	36931977	C/T	0.49	0.52	0.317
rs1951771	61961	36932361	A/G	0.93	NA	NA
rs2376323	63351	36933751	G/T			
rs2013358	63926	36934326	A/G	0.13	0.13	0.821
rs2181494	65798	36936198	A/G	0.42	0.43	0.512
rs1957728	66043	36936443	A/C			
rs1957729	66044	36936444	A/G	0.79	0.77	0.405
rs1957730	66246	36936646	C/T	0.15	0.16	0.719
rs1957731	66318	36936718	C/T	0.14	0.16	0.413

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1998468	66547	36936947	G/T	0.13	0.12	0.468
rs1957732	71238	36941638	C/T	0.10	0.10	0.841
rs1957733	71283	36941683	A/G	0.63	0.61	0.632
rs2376322	71492	36941892	A/G	0.26	0.28	0.509
rs2889345	72274	36942674	A/G	0.20	0.18	0.234
rs1815267	73762	36944162	A/T	0.46	0.45	0.674
rs1957734	74209	36944609	G/T	0.55	0.64	0.003
rs1957735	75284	36945684	A/T	0.63	0.61	0.430
rs1957736	77347	36947747	A/C	0.05	0.05	0.903
rs1957737	77589	36947989	C/T	0.71	0.75	0.164
rs1957738	78096	36948496	A/G			
rs1957739	78606	36949006	A/G			
rs1957740	78862	36949262	G/T			
rs1957741	79135	36949535	A/G	0.76	0.80	0.077
rs1957742	79146	36949546	A/G	0.95	0.96	0.500
rs1957743	79456	36949856	C/T	0.21	0.16	0.039
rs1957744	79609	36950009	A/G	0.66	0.70	0.088
rs1957745	80086	36950486	A/G	0.88	0.90	0.354
rs1957746	80119	36950519	A/G	0.40	0.44	0.120
rs1957747	80766	36951166	C/T	0.72	0.76	0.093
rs2146670	81110	36951510	A/G	0.73	0.77	0.072
rs2146671	81269	36951669	A/T	0.17	0.15	0.250
rs1957748	81668	36952068	C/T	0.16	0.14	0.407
rs2162307	82433	36952833	C/T	0.73	0.76	0.170
rs1962839	82559	36952959	C/G			
rs2376315	83298	36953698	C/T	0.62	0.66	0.179
rs1426410	83821	36954221	A/G	0.75	0.77	0.307
rs1895921	84121	36954521	C/T	0.75	0.78	0.175
rs1895922	84147	36954547	C/T	0.15	0.12	0.095
rs1035779	84543	36954943	A/G	0.66	0.64	0.649
rs1035780	84554	36954954	A/G			
rs1035781	84691	36955091	A/G	0.73	0.77	0.100
rs1035782	84727	36955127	A/G			
rs1426411	85678	36956078	C/T	0.76	0.80	0.084
rs1834602	86699	36957099	C/T	0.20	0.16	0.072
rs1834603	86700	36957100	A/G	0.94	0.92	0.326
rs1834604	86792	36957192	A/G	0.70	0.73	0.287
rs1834605	86832	36957232	A/G	0.72	0.76	0.057
rs2162308	87045	36957445	A/G			
rs1365341	87140	36957540	A/G	0.18	0.15	0.086
rs1820458	87365	36957765	A/C	0.23	0.21	0.298
rs1469310	88342	36958742	C/T	0.20	0.18	0.265
rs3057879	88498	36958898	-T/C	0.70	0.71	0.649
rs1469311	88589	36958989	A/G	0.70	0.74	0.065
rs768326	95502	36965902	A/G			
rs1863523	96968	36967368	C/T	0.21	0.18	0.247
rs1469312	97448	36967848	C/T	0.78	0.76	0.312
rs1469313	97568	36967968	C/T	0.81	0.80	0.617
rs1951773	98724	36969124	C/T			
rs2120655	Not mapped	Not mapped	T/G			
rs2181495	Not mapped	Not mapped	G/A	0.78	0.76	0.617

[0236] The *chrom 4* proximal SNPs were also allelotyped in the replication cohorts using the methods described herein and the primers provided in Tables 15 and 16. The replication allelotyping results for replication cohort #1 and replication cohort #2 are provided in Tables 18 and 19, respectively.

**TABLE 18**

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs3849023	211	36870611	G/T			
rs1444311	7217	36877617	A/G	0.27	0.25	0.441
rs2044295	7895	36878295	A/C			
rs2166093	13308	36883708	C/T			
rs2376334	14279	36884679	G/T	0.14	0.14	0.970
rs1444320	17026	36887426	C/T			
rs2044294	18271	36888671	A/G	0.16	0.14	0.423
rs1899864	20417	36890817	C/T			
rs1562094	21843	36892243	A/G	0.22	0.21	0.725
rs1562098	22069	36892469	A/G			
rs1562097	22145	36892545	A/G			
rs1562096	22519	36892919	A/G	0.20	0.19	0.795
rs1562095	22539	36892939	A/G	0.55	0.53	0.453
rs1444319	23236	36893636	A/C	0.70	0.80	0.003
rs1444318	23256	36893656	A/C	0.12	0.13	0.645
rs1025938	23402	36893802	C/T	0.18	0.18	0.824
rs1025937	23499	36893899	A/C			
rs1025936	23620	36894020	C/T	0.85	0.83	0.622
rs1020333	23871	36894271	A/T			
rs2120654	24136	36894536	C/G	0.16	0.16	0.914
rs2588547	25427	36895827	A/G	0.40	0.40	0.980
rs2044293	25866	36896266	G/T			
rs2760324	26541	36896941	A/G	0.57	0.61	0.287
rs2588546	26576	36896976	G/T	0.08	0.05	0.265
rs2588545	26689	36897089	A/G			
rs2760328	26720	36897120	A/C	0.25	untyped	NA
rs2588544	27113	36897513	C/T			
rs2760331	27164	36897564	C/T	0.88	0.92	0.193
rs2588543	27186	36897586	A/G	0.57	0.58	0.869
rs2588542	28341	36898741	A/T			
rs2588541	29160	36899560	C/T	0.61	0.57	0.230
rs2588540	29844	36900244	A/G	0.64	0.64	0.926
rs2760336	30665	36901065	G/T			
rs2760337	30830	36901230	A/G	0.16	0.16	0.956
rs2028732	31061	36901461	A/C	0.60	0.57	0.330
rs2588538	31523	36901923	C/T	0.62	0.61	0.747
rs1992617	32326	36902726	C/T	0.62	0.59	0.341
rs1998469	32346	36902746	A/G			
rs1998470	32358	36902758	C/T	0.78	0.88	-0.0001
rs1975498	34909	36905309	C/T			
rs1562093	34975	36905375	A/G	0.89	0.90	0.905
rs1975497	35066	36905466	C/T	0.12	0.12	0.873
rs1562092	35096	36905496	G/T			
rs2248788	35375	36905775	C/T	0.28	0.31	0.308
rs1899862	36304	36906704	A/G	0.19	0.14	0.088
rs2588532	36712	36907112	A/T	0.30	0.33	0.347
rs1885878	36770	36907170	C/T	0.36	0.34	0.362
rs986648	37342	36907742	C/T	0.74	0.75	0.773
rs986647	37412	36907812	C/T	0.78	0.77	0.693

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs1010010	37884	36908284	A/G	0.25	0.26	0.690
rs1010009	38077	36908477	A/C	0.27	0.26	0.870
rs2760325	38300	36908700	C/T			
rs2588531	38301	36908701	C/T			
rs1838388	41189	36911589	C/T	0.74	0.75	0.826
rs1975495	44408	36914808	C/T			
rs2181491	44493	36914893	A/C	0.16	0.10	0.057
rs1975496	44571	36914971	A/G	0.25	0.26	0.596
rs2181492	44670	36915070	A/G	0.11	0.08	0.167
rs2224719	45219	36915619	A/G	0.78	0.79	0.705
rs2224720	45258	36915658	C/T	0.19	0.21	0.478
rs1951770	47261	36917661	A/G	0.25	0.16	0.003
rs2296040	48473	36918873	A/C	0.39	0.43	0.241
rs1957723	48771	36919171	A/G	0.44	0.36	0.027
rs1957725	55292	36925692	C/T	0.73	0.80	0.024
rs2889346	56479	36926879	A/G	0.53	0.55	0.552
rs1885879	56747	36927147	A/C	0.43	0.50	0.057
rs1957726	60620	36931020	G/T	0.14	0.14	0.918
rs1957727	60688	36931088	A/C	0.71	0.78	0.038
rs1885880	61058	36931458	A/C	0.44	0.42	0.627
rs1885881	61129	36931529	C/T	0.12	0.12	0.833
rs942108	61577	36931977	C/T	0.42	0.49	0.096
rs1951771	61961	36932361	A/G	0.93	NA	NA
rs2376323	63351	36933751	G/T			
rs2013358	63926	36934326	A/G	0.13	0.12	0.795
rs2181494	65798	36936198	A/G	0.38	0.41	0.424
rs1957728	66043	36936443	A/C			
rs1957729	66044	36936444	A/G	0.78	0.77	0.672
rs1957730	66246	36936646	C/T	0.15	0.15	0.885
rs1957731	66318	36936718	C/T	0.15	0.16	0.719
rs1998468	66547	36936947	G/T	0.14	0.10	0.243
rs1957732	71238	36941638	C/T	0.10	0.09	0.817
rs1957733	71283	36941683	A/G	0.62	NA	0.628
rs2376322	71492	36941892	A/G	0.26	0.27	0.660
rs2889345	72274	36942674	A/G	0.22	0.16	0.020
rs1815267	73762	36944162	A/T	0.46	0.48	0.626
rs1957734	74209	36944609	G/T	0.44	0.61	~0.0001
rs1957735	75284	36945684	A/T	0.63	0.63	0.792
rs1957736	77347	36947747	A/C	0.03	0.03	0.987
rs1957737	77589	36947989	C/T	0.69	0.77	0.024
rs1957738	78096	36948496	A/G			
rs1957739	78606	36949006	A/G			
rs1957740	78862	36949262	G/T			
rs1957741	79135	36949535	A/G	0.75	0.83	0.008
rs1957742	79146	36949546	A/G	0.94	0.96	0.459
rs1957743	79456	36949856	C/T	0.24	0.14	0.009
rs1957744	79609	36950009	A/G	0.63	0.72	0.006
rs1957745	80086	36950486	A/G	0.86	0.90	0.229
rs1957746	80119	36950519	A/G	0.42	0.50	0.019
rs1957747	80766	36951166	C/T	0.71	0.79	0.009
rs2146670	81110	36951510	A/G	0.72	0.81	0.004
rs2146671	81269	36951669	A/T	0.17	0.13	0.106
rs1957748	81668	36952068	C/T	0.17	0.13	0.133
rs2162307	82433	36952833	C/T	0.72	0.78	0.020
rs1962839	82559	36952959	C/G			
rs2376315	83298	36953698	C/T	0.61	0.67	0.074
rs1426410	83821	36954221	A/G	0.73	0.79	0.058



dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs1895921	84121	36954521	C/T	0.72	0.80	0.013
rs1895922	84147	36954547	C/T	0.17	0.11	0.014
rs1035779	84543	36954943	A/G	0.66	0.64	0.613
rs1035780	84554	36954954	A/G			
rs1035781	84691	36955091	A/G	0.71	0.78	0.059
rs1035782	84727	36955127	A/G			
rs1426441	85678	36956078	C/T	0.75	0.82	0.008
rs1834602	86699	36957099	C/T	0.22	0.15	0.020
rs1834603	86700	36957100	A/G	0.94	0.92	0.483
rs1834604	86792	36957192	A/G	0.69	0.75	0.056
rs1834605	86832	36957232	A/G	0.71	0.79	0.007
rs2162308	87045	36957445	A/G			
rs1365341	87140	36957540	A/G	0.19	0.13	0.017
rs1820458	87365	36957765	A/C	0.24	0.19	0.141
rs1469310	88342	36958742	C/T	0.22	0.17	0.061
rs3057879	88498	36958898	-T/C	0.67	NA	NA
rs1469311	88589	36958989	A/G	0.67	0.76	0.006
rs768326	95502	36965902	A/G			
rs1863523	96968	36967368	C/T	0.22	0.17	0.103
rs1469312	97448	36967848	C/T	0.80	0.77	0.236
rs1469313	97568	36967968	C/T	0.83	0.80	0.422
rs1951773	98724	36969124	C/T			
rs2120655	Not mapped	Not mapped	T/G			
rs2181495	Not mapped	Not mapped	G/A	0.78	0.76	0.617

TABLE 19

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs3849023	211	36870611	G/T			
rs1444311	7217	36877617	A/G	0.25	0.26	0.876
rs2044295	7895	36878295	A/C			
rs2166093	13308	36883708	C/T			
rs2376334	14279	36884679	G/T	0.16	0.18	0.532
rs1444320	17026	36887426	C/T			
rs2044294	18271	36888671	A/G	NA	0.15	NA
rs1899864	20417	36890817	C/T			
rs1562094	21843	36892243	A/G	NA	0.28	NA
rs1562098	22069	36892469	A/G			
rs1562097	22145	36892545	A/G			
rs1562096	22519	36892919	A/G	0.20	0.23	0.364
rs1562095	22539	36892939	A/G	0.50	0.48	0.588
rs1444319	23236	36893636	A/C	0.79	0.79	0.923
rs1444318	23256	36893656	A/C	0.12	0.13	0.711
rs1025938	23402	36893802	C/T	0.18	0.22	0.247
rs1025937	23499	36893899	A/C			
rs1025936	23620	36894020	C/T	0.84	0.86	0.403
rs1020333	23871	36894271	A/T			
rs2120654	24136	36894536	C/G	0.14	0.16	0.682
rs2588547	25427	36895827	A/G	0.37	NA	
rs2044293	25866	36896266	G/T			
rs2760324	26541	36896941	A/G	0.60	0.60	0.965
rs2588546	26576	36896976	G/T	0.05	0.06	0.797
rs2588545	26689	36897089	A/G			
rs2760328	26720	36897120	A/C	0.25	0.26	0.816

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs2588544	27113	36897513	C/T			
rs2760331	27164	36897564	C/T	0.95	0.96	0.597
rs2588543	27186	36897586	A/G	0.60	0.61	0.801
rs2588542	28341	36898741	A/T			
rs2588541	29160	36899560	C/T	0.62	0.61	0.972
rs2588540	29844	36900244	A/G	0.60	0.59	0.810
rs2760336	30665	36901065	G/T			
rs2760337	30830	36901230	A/G	0.16	0.17	0.659
rs2028732	31061	36901461	A/C	0.60	0.60	0.976
rs2588538	31523	36901923	C/T	0.61	0.61	0.912
rs1992617	32326	36902726	C/T	0.61	0.59	0.583
rs1998469	32346	36902746	A/G			
rs1998470	32358	36902758	C/T	0.84	0.81	0.338
rs1975498	34909	36905309	C/T			
rs1562093	34975	36905375	A/G	0.88	0.84	0.199
rs1975497	35066	36905466	C/T	0.13	0.15	0.613
rs1562092	35096	36905496	G/T			
rs2248788	35375	36905775	C/T	0.30	0.31	0.884
rs1899862	36304	36906704	A/G	0.17	0.18	0.563
rs2588532	36712	36907112	A/T	0.29	0.29	0.952
rs1885878	36770	36907170	C/T	0.33	0.36	0.405
rs986648	37342	36907742	C/T	0.75	0.72	0.283
rs986647	37412	36907812	C/T	0.79	0.74	0.186
rs1010010	37884	36908284	A/G	0.25	0.26	0.843
rs1010009	38077	36908477	A/C	0.25	0.24	0.764
rs2760325	38300	36908700	C/T			
rs2588531	38301	36908701	C/T			
rs1838388	41189	36911589	C/T	0.76	0.72	0.284
rs1975495	44408	36914808	C/T			
rs2181491	44493	36914893	A/C	0.12	0.15	0.464
rs1975496	44571	36914971	A/G	0.27	0.25	0.577
rs2181492	44670	36915070	A/G	0.10	0.11	0.844
rs2224719	45219	36915619	A/G	0.78	0.75	0.426
rs2224720	45258	36915658	C/T	0.21	0.20	0.790
rs1951770	47261	36917661	A/G	0.19	0.20	0.796
rs2296040	48473	36918873	A/C	0.43	0.42	0.804
rs1957723	48771	36919171	A/G	0.41	0.42	0.653
rs1957725	55292	36925692	C/T	0.77	0.75	0.439
rs2889346	56479	36926879	A/G	0.56	0.56	0.948
rs1885879	56747	36927147	A/C	0.46	0.45	0.959
rs1957726	60620	36931020	G/T	0.14	0.15	0.673
rs1957727	60688	36931088	A/C	0.76	0.73	0.255
rs1885880	61058	36931458	A/C	0.43	0.45	0.614
rs1885881	61129	36931529	C/T	0.13	0.10	0.346
rs942108	61577	36931977	C/T	0.58	0.56	0.730
rs1951771	61961	36932361	A/G			
rs2376323	63351	36933751	G/T			
rs2013358	63926	36934326	A/G	0.13	0.15	0.469
rs2181494	65798	36936198	A/G	0.46	0.47	0.820
rs1957728	66043	36936443	A/C			
rs1957729	66044	36936444	A/G	0.80	0.78	0.440
rs1957730	66246	36936646	C/T	0.15	0.17	0.668
rs1957731	66318	36936718	C/T	0.14	0.16	0.387
rs1998468	66547	36936947	G/T	0.12	0.13	0.615
rs1957732	71238	36941638	C/T	0.09	0.11	0.469
rs1957733	71283	36941683	A/G	0.60	0.02	
rs2376322	71492	36941892	A/G	0.27	0.29	0.582

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs2889345	72274	36942674	A/G	0.17	0.21	0.308
rs1815267	73762	36944162	A/T	0.46	0.41	0.151
rs1957734	74209	36944609	G/T	0.68	0.69	0.766
rs1957735	75284	36945684	A/T	0.62	0.58	0.311
rs1957736	77347	36947747	A/C	0.07	0.08	0.688
rs1957737	77589	36947989	C/T	0.75	0.71	0.305
rs1957738	78096	36948496	A/G			
rs1957739	78606	36949006	A/G			
rs1957740	78862	36949262	G/T			
rs1957741	79135	36949535	A/G	0.78	0.76	0.446
rs1957742	79146	36949546	A/G	0.96	0.96	0.938
rs1957743	79456	36949856	C/T	0.17	0.19	0.667
rs1957744	79609	36950009	A/G	0.69	0.66	0.423
rs1957745	80086	36950486	A/G	0.90	0.89	0.738
rs1957746	80119	36950519	A/G	0.37	0.35	0.708
rs1957747	80766	36951166	C/T	0.72	0.70	0.639
rs2146670	81110	36951510	A/G	0.75	0.72	0.306
rs2146671	81269	36951669	A/T	0.16	0.17	0.806
rs1957748	81668	36952068	C/T	0.14	0.17	0.453
rs2162307	82433	36952833	C/T	0.76	0.73	0.465
rs1962839	82559	36952959	C/G			
rs2376315	83298	36953698	C/T	0.64	0.63	0.767
rs1426410	83821	36954221	A/G	0.77	0.74	0.465
rs1895921	84121	36954521	C/T	0.78	0.75	0.320
rs1895922	84147	36954547	C/T	0.12	0.13	0.586
rs1035779	84543	36954943	A/G	NA	0.65	NA
rs1035780	84554	36954954	A/G			
rs1035781	84691	36955091	A/G	0.75	0.76	0.830
rs1035782	84727	36955127	A/G			
rs1426411	85678	36956078	C/T	0.78	0.76	0.488
rs1834602	86699	36957099	C/T	0.18	0.18	0.945
rs1834603	86700	36957100	A/G	0.95	NA	
rs1834604	86792	36957192	A/G	0.72	0.69	0.427
rs1834605	86832	36957232	A/G	0.73	0.72	0.647
rs2162308	87045	36957445	A/G			
rs1365341	87140	36957540	A/G	0.16	0.17	0.667
rs1820458	87365	36957765	A/C	0.23	0.24	0.670
rs1469310	88342	36958742	C/T	0.19	0.21	0.592
rs3057879	88498	36958898	-T/C	0.74	0.71	0.478
rs1469311	88589	36958989	A/G	0.74	0.72	0.582
rs769326	95502	36965902	A/G			
rs1863523	96968	36967368	C/T	0.20	0.21	0.687
rs1469312	97448	36967848	C/T	0.76	0.75	0.807
rs1469313	97568	36967968	C/T	0.78	0.79	0.824
rs1951773	98724	36969124	C/T			
rs2120655	Not mapped	Not mapped	T/G			
rs2181495	Not mapped	Not mapped	G/A			

[0237] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1B for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the

estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1B can be determined by consulting Table 17. For example, the left-most X on the left graph is at position 36870611. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0238] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than  $10^{-8}$  were truncated at that value.

[0239] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

#### Example 6

##### SNW1 Region Proximal SNPs

[0240] SNP rs1477261 is associated with osteoarthritis and is described in Table A. It lies within an intron of the SKI-interacting protein gene (SNW1). This gene, a member of the SNW gene family, encodes a coactivator that enhances transcription from some Pol II promoters. This coactivator can bind to the ligand-binding domain of the vitamin D receptor and to retinoid receptors to enhance vitamin D-, retinoic acid-, estrogen-, and glucocorticoid-mediated gene expression. It also can interact with poly(A)-binding protein 2 to directly control the expression of muscle-specific genes at the transcriptional level. One hundred sixty-three additional allelic variants proximal to rs1477261 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 20. The chromosome position provided in column four of Table 20 is based on Genome "Build 34" of NCBI's GenBank.

TABLE 20

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs7143926	14	218	76161268	a/t
rs1549071	14	1440	76162490	c/t
rs8012858	14	1442	76162492	c/t
rs7155611	14	2611	76163661	c/t
rs176941	14	4317	76165367	a/c
rs176942	14	4724	76165774	a/g
rs176943	14	4788	76165838	g/t
rs176944	14	5202	76166252	g/t
rs4365221	14	5780	76166830	c/t
rs3168952	14	5974	76167024	c/t
rs176945	14	6644	76167694	c/g
rs176946	14	7430	76168480	a/g
rs176947	14	7938	76168988	c/t
rs176948	14	8095	76169145	c/t
rs176949	14	8183	76169233	a/c
rs176950	14	8312	76169362	c/t
rs176951	14	8352	76169402	a/c
rs7156905	14	9348	76170398	c/t
rs3217197	14	9378	76170428	-/tctc
rs2270443	14	9617	76170667	a/g
rs176952	14	9727	76170777	c/t
rs176953	14	9834	76170884	c/t
rs176954	14	9899	76170949	g/t
rs176955	14	10211	76171261	c/t
rs3214416	14	10377	76171427	-/t
rs176956	14	10695	76171745	c/t
rs2544566	14	10729	76171779	c/g
rs2544567	14	10730	76171780	c/t
rs176957	14	11433	76172483	a/g
rs176958	14	11951	76173001	c/g
rs176959	14	12697	76173747	c/t
rs1802227	14	12982	76174032	a/c
rs176961	14	14419	76175469	c/t
rs176962	14	14501	76175551	c/t
rs7401285	14	14983	76176033	a/c
rs176963	14	15280	76176330	c/t
rs176964	14	15475	76176525	a/g
rs4903631	14	15888	76176938	a/g
rs4903632	14	15976	76177026	a/t
rs176965	14	16307	76177357	a/c
rs4903633	14	16442	76177492	a/c
rs176966	14	17255	76178305	c/t
rs176968	14	18948	76179998	g/t
rs176969	14	19435	76180485	a/t
rs176970	14	19753	76180803	c/t
rs7149198	14	20021	76181071	c/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs7147918	14	20022	76181072	a/c
rs7148685	14	20503	76181553	a/g
rs1184232	14	20590	76181640	g/t
rs1184233	14	21804	76182854	g/t
rs1184234	14	21919	76182969	c/t
rs7401998	14	21990	76183040	a/t
rs176974	14	22412	76183462	a/g
rs6574390	14	22536	76183586	c/t
rs176975	14	23432	76184482	a/g
rs176976	14	23468	76184518	g/t
rs176977	14	23772	76184822	c/t
rs8013727	14	24325	76185375	c/t
rs176978	14	24773	76185823	c/t
rs2111829	14	26274	76187324	c/t
rs176980	14	27440	76188490	c/g
rs5809848	14	28561	76189611	-/acag
rs5809849	14	30071	76191121	-/a
rs4383070	14	31764	76192814	a/t
rs7493652	14	33008	76194058	c/t
rs2112133	14	35310	76196360	a/t
rs1963833	14	35460	76196510	a/c
rs6574391	14	37112	76198162	a/g
rs7155062	14	37285	76198335	a/g
rs4899674	14	37747	76198797	c/t
rs8022516	14	38057	76199107	c/t
rs7140838	14	38859	76199909	a/c
rs7141127	14	38860	76199910	a/g
rs6574392	14	39525	76200575	a/g
rs8003691	14	40216	76201266	a/g
rs8003979	14	40281	76201331	c/t
rs8010541	14	41453	76202503	c/g
rs8016416	14	42091	76203141	a/t
rs8016175	14	42513	76203563	a/g
rs7154571	14	42935	76203985	c/t
rs7158826	14	42985	76204035	a/g
rs7159310	14	43003	76204053	a/g
rs7401900	14	43281	76204331	a/g
rs7160355	14	43716	76204766	c/t
rs2032781	14	43866	76204916	a/g
rs6574394	14	44234	76205284	g/t
rs8007598	14	44596	76205646	a/g
rs2267767	14	44871	76205921	c/t
rs6574395	14	45005	76206055	a/g
rs7150066	14	45282	76206332	a/c
rs7492334	14	47178	76208228	a/c
rs4359361	14	47816	76208866	g/t
rs4605089	14	47887	76208937	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs7146446	14	48134	76209184	c/t
rs4346144	14	48135	76209185	a/g
rs7148078	14	48276	76209326	g/t
rs7148286	14	48400	76209450	c/t
rs3783980	14	48798	76209848	a/g
rs1549119	14	48803	76209853	a/t
rs1984925	14	49146	76210196	c/t
rs1477261	14	49969	76211019	a/t
rs8016447	14	51059	76212109	a/g
rs7494044	14	51064	76212114	c/t
rs2023288	14	53285	76214335	a/t
rs7151685	14	54560	76215610	c/t
rs2112135	14	54748	76215798	a/g
rs2161088	14	54785	76215835	c/g
rs4903638	14	55102	76216152	c/g
rs1477262	14	55644	76216694	a/g
rs1477263	14	55705	76216755	g/t
rs1477264	14	55841	76216891	a/g
rs2277917	14	56623	76217673	c/g
rs2277918	14	56825	76217875	a/c
rs2277919	14	56827	76217877	a/g
rs1978416	14	56892	76217942	c/t
rs3759728	14	59150	76220200	a/t
rs6574399	14	59958	76221008	a/t
rs7155336	14	60231	76221281	c/t
rs7156186	14	60524	76221574	a/g
rs7142390	14	61871	76222921	c/t
rs7145875	14	62226	76223276	c/t
rs8014635	14	63230	76224280	g/t
rs8015938	14	63468	76224518	g/t
rs8015313	14	63787	76224837	c/t
rs8006315	14	65732	76226782	a/c
rs6574400	14	65989	76227039	a/g
rs7140816	14	68832	76229882	g/t
rs4566078	14	69904	76230954	c/t
rs7141050	14	70365	76231415	a/g
rs3049356	14	70886	76231936	-/tatc
rs4903639	14	73088	76234138	a/t
rs4903641	14	73103	76234153	c/t
rs2364838	14	75934	76236984	c/t
rs2364839	14	75966	76237016	c/t
rs4632066	14	76273	76237323	c/t
rs2112136	14	77943	76238993	c/t
rs4641655	14	78466	76239516	c/t
rs4635269	14	78861	76239911	c/t
rs4570764	14	78872	76239922	a/g
rs759808	14	79836	76240886	g/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs7150531	14	80908	76241958	c/t
rs7154968	14	81509	76242559	c/g
rs7146657	14	83576	76244626	c/t
rs7145859	14	83662	76244712	c/g
rs4903643	14	83782	76244832	c/t
rs717682	14	84282	76245332	g/t
rs717683	14	84444	76245494	a/g
rs1477259	14	85129	76246179	c/g
rs8019064	14	85151	76246201	a/g
rs8018971	14	85296	76246346	a/c
rs1477260	14	85809	76246859	c/g
rs5809851	14	86387	76247437	-/t
rs1985149	14	86494	76247544	a/g
rs1008988	14	89786	76250836	a/g
rs1008989	14	89894	76250944	a/t
rs8018222	14	90122	76251172	g/t
rs1006040	14	92067	76253117	a/g
rs1006039	14	92187	76253237	c/t
rs1006038	14	92312	76253362	a/g
rs8009784	14	92824	76253874	g/t
rs4903644	14	93733	76254783	c/t
rs7149496	14	96553	76257603	c/g
rs6574402	14	96941	76257991	a/c

#### Assay for Verifying and Allelotyping SNPs

[0241] The methods used to verify and allelotype the 101 proximal SNPs of Table 20 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 21 and Table 22, respectively.

TABLE 21

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs7143926	ACGTTGGATGGAGTCACCCAAAATTAAGGC	ACGTTGGATGGAAAGCCAAAATTAGCCTGC
rs1549071	ACGTTGGATGGTGAGACGCTGTCTCAGTAA	ACGTTGGATGCTCCACACTTGGAGAAGTTG
rs8012858	ACGTTGGATGGTGAGACGCTGTCTCAGTAA	ACGTTGGATGCTCCACACTTGGAGAAGTTG
rs7155611	ACGTTGGATGATGGAATACAGGCACCGTTC	ACGTTGGATGCCCTTCTTAATCTCCATGG
rs176941	ACGTTGGATGTTAGTATGGGAAAAGGGCTC	ACGTTGGATGCAACAATCCTATGAGTTGGG
rs176942	ACGTTGGATGAGTGGCTCAGATGTGAGTAG	ACGTTGGATGTGGCTTCCACCAACCACATG
rs176943	ACGTTGGATGACCAAGCCCAAGTAAAGTCTC	ACGTTGGATGGCATCCGCAAGATGCTAATG
rs176944	ACGTTGGATGGGCGCTCAATATTGGCTAAATG	ACGTTGGATGCTTAACCATTAAGAGCCCTTC
rs4365221	ACGTTGGATGAAATAAGGCAGGAAGGGTAG	ACGTTGGATGTCCCAACTTACTGGTCTTTC
rs3168952	ACGTTGGATGATGTACCAGACTTGGTGGTG	ACGTTGGATGTTTGCTGAGGATGGAGACTG
rs176945	ACGTTGGATGCCTACTATACACTCACAAAA	ACGTTGGATGTTTTTAAACACACTTAAAGC



dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs176946	ACGTTGGATGGCTTTATCATAGGATTTGTG	ACGTTGGATGGAGAGATGTGTGTTTGTGAG
rs176947	ACGTTGGATGTGAGTAGCTGGGACTACAGG	ACGTTGGATGGGCCAACATAGCGAACTCC
rs176948	ACGTTGGATGCGAGAGCCAAAGGTCAACAAG	ACGTTGGATGTACAGGTGTGAGCCCTCATG
rs176949	ACGTTGGATGTAGGAACCTCCCTGCAGTTCC	ACGTTGGATGCCCTTGCTGGCTTTAAAGAG
rs176950	ACGTTGGATGAATCAGGAGTGACATCCC	ACGTTGGATGTGGAGGAGAAACCTGACITG
rs176951	ACGTTGGATGCCCTATATAATCTCCTCCCC	ACGTTGGATGCGAGAGTGACATCCCATTAC
rs1756905	ACGTTGGATGTGAGAGAGAGAACCTGTCTC	ACGTTGGATGAAAGCGCGCTTTGATGTTGG
rs3217197	ACGTTGGATGTTGATTGTGCCACTGCACCTC	ACGTTGGATGACTCTAGTTGGAATCCTGG
rs2270443	ACGTTGGATGATAACTCAGTCCAGGTGTGG	ACGTTGGATGCACTCAAGCAGCTACTCAC
rs176952	ACGTTGGATGGATCTCAGCTCACTGCAATC	ACGTTGGATGTATCTGGGTGACTGAGGAAG
rs176953	ACGTTGGATGTTGAGGTCAAGGATTTGGGA	ACGTTGGATGGCCACCACACCCAGCTAATT
rs176954	ACGTTGGATGAAACATAGGCCAGGTGCAG	ACGTTGGATGAAACTCCTGACCTCAAGCCA
rs176955	ACGTTGGATGCTAGAGTGCTTGATGTACC	ACGTTGGATGGTCACTCAGGAGCTAGAG
rs3214416	ACGTTGGATGACGACTATCATCAGCTGTTCT	ACGTTGGATGACCAAGAAGCTGTAAGCTAG
rs176956	ACGTTGGATGTACAGGCATAAGCCACCATG	ACGTTGGATGAGGAAGGGGTGTAACCAAGG
rs2544566	ACGTTGGATGCAAGCAATCTTCCCATTCTGG	ACGTTGGATGTGATCCGATTTTGGCTGGG
rs2544567	ACGTTGGATGCAAGCAATCTTCCCATTCTGG	ACGTTGGATGTGATCCGATTTTGGCTGGG
rs176957	ACGTTGGATGTTTACCCTGTTAGCCAGGA	ACGTTGGATGTAATCCGAGCATTTGGGAG
rs176958	ACGTTGGATGAAACTGGGCACTCTACCAC	ACGTTGGATGAAATCGGCCATTGCACTC
rs176959	ACGTTGGATGCAAGGCAGTTTTTATTTGTCCC	ACGTTGGATGGGTAGGGAGTCAATAACC
rs1802227	ACGTTGGATGAACAATAGTTGCACCAAG	ACGTTGGATGTTTTAATTTGGAGTGGGCA
rs176961	ACGTTGGATGAACCCAGTTTAAGACCGGCC	ACGTTGGATGTACAGGTGTGTGCCACCATTG
rs176962	ACGTTGGATGATATTTCTGGCTGGGCACTG	ACGTTGGATGACTGGGTTCAAGCAATCTGC
rs7401285	ACGTTGGATGACAGAGTGGGACTCCATATC	ACGTTGGATGGTCTCAAAGTGGGTGCTCTG
rs176963	ACGTTGGATGTAAAGCCTGGGAAAAACACAG	ACGTTGGATGCCCACTCTACTTTCCAGTAG
rs176964	ACGTTGGATGAGAGTCAGTGTCTACAAAA	ACGTTGGATGTAATCCCGCTTTACAGCTTC
rs4903631	ACGTTGGATGGTAAATGCCAGCATGATGAC	ACGTTGGATGCTCAGCCCACTATAAAGAAG
rs4903632	ACGTTGGATGTGTGAATACCTATCCTCAGG	ACGTTGGATGGTCTCATGCTCAGGCAATTAC
rs176965	ACGTTGGATGAATGCTTTATAAGGGCTGCC	ACGTTGGATGCTCAGAAAACAAAGGATGTG
rs4903633	ACGTTGGATGCAACCCCCAACCATCATAT	ACGTTGGATGCTAACAGATTGTTGACATGG
rs176966	ACGTTGGATGCTCTCGAGTAGCTGGGACTA	ACGTTGGATGGTGGCCAAACATGGTGAACCCG
rs176968	ACGTTGGATGGCGAAACTCCGCTCTCAAAC	ACGTTGGATGTAGTGATCTTCCCACTTAGC
rs176969	ACGTTGGATGCTGTCTGTCCGATTACTGC	ACGTTGGATGTCTAGAATCAAGCATGCGGC
rs176970	ACGTTGGATGCTAATGTCTCTAGTACAGTGG	ACGTTGGATGCTTCTCTAGCTATTTTGC
rs7149198	ACGTTGGATGCAATGGGATATTACTACGCC	ACGTTGGATGTTTCTGTGCCGGGCTTATTC
rs7147918	ACGTTGGATGCAATGGGATATTACTACGCC	ACGTTGGATGTTTCTGTGCCGGGCTTATTC
rs7148685	ACGTTGGATGTGTCTCTTTTGAGACCGCTC	ACGTTGGATGCTCAATCGCAAGAAACAGAG
rs1184232	ACGTTGGATGAGAGGCGACCTACAGAATG	ACGTTGGATGCTCTGTTCTTTTGGCATGAG
rs1184233	ACGTTGGATGAAGTGTGTGGATTACAGGTG	ACGTTGGATGAGTGAAGATCGCCACAAG
rs1184234	ACGTTGGATGGCTATGTGCAGTGACTCATG	ACGTTGGATGCTCAGACCTCAGGTGATCT
rs17401998	ACGTTGGATGTGAGTAGTACAGGACAACAGG	ACGTTGGATGTAAGCTGGTGAACCCCATGCT
rs176974	ACGTTGGATGTTACAGCGAGCTGAGATCAT	ACGTTGGATGAGGATCACTACTGCTTCAG
rs6574390	ACGTTGGATGATGATAAACCCCGTCTGTAC	ACGTTGGATGCCCTGAGTAGCTGGGATTAC
rs176975	ACGTTGGATGTGTAGAATCTAGGTGGTAGG	ACGTTGGATGCCAGCCCTTCTGTGACATTTT
rs176976	ACGTTGGATGGGTAGGAGATACAGGTGTTCT	ACGTTGGATGCCAGCCCTTCTGTGACATTTT
rs176977	ACGTTGGATGTTGCATCTTACACTTCAGC	ACGTTGGATGGGGAAACATATGCTGAATCTC
rs8013727	ACGTTGGATGTGCCCTGGTTGTATACCTAAC	ACGTTGGATGCTTGAGAACGATTCTGTTGTC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs176978	ACGTTGGATGGGGACCATGTTTTGTACC	ACGTTGGATGAATCTGTGGAATGGGCATG
rs2111829	ACGTTGGATGCATGTGGAAAAGGTATGAC	ACGTTGGATGCCCTACTTTATATGCAGTAGG
rs176980	ACGTTGGATGATGGCCAATGCTATGAACGC	ACGTTGGATGAAGGGCAGTTGCAGGAAAAG
rs5809848	ACGTTGGATGTCATTTTTCCAGAGCTTGGG	ACGTTGGATGCCATTTTCTACGTGCTTTGG
rs5809849	ACGTTGGATGGTGAATACCGTGTCAGTTCC	ACGTTGGATGTGCAGTGAGCTGAGATCATG
rs4383070	ACGTTGGATGAGCGATTCTCTTGCTCAGC	ACGTTGGATGAACCTAGCTGGGCATTGTGG
rs7493652	ACGTTGGATGGGTCATATACCACAAGTAAC	ACGTTGGATGCTGGCCCTATGCTATTTTCA
rs212133	ACGTTGGATGGCCACCACAACCTGGCTAATT	ACGTTGGATGTGTGGTCAGGAGATCGAAAC
rs1963833	ACGTTGGATGTAAGCCAAGATTGCGCTCACT	ACGTTGGATGAGCATTAAAGGTAGAATGCC
rs6574391	ACGTTGGATGTAACCGTTGCTATGGAGAAG	ACGTTGGATGACCTATACCAACCCTAAGCTG
rs7155062	ACGTTGGATGGCTCCTTATTTGGGCATTCC	ACGTTGGATGCACCTCAGCCTTGTGAGATAC
rs4899674	ACGTTGGATGAATGTGCTGAGGAACTGAG	ACGTTGGATGGCTCTGTGACTTTCAAGAG
rs8022516	ACGTTGGATGGTTTGAAGGCATCTTTTGGG	ACGTTGGATGCTAGCCTGGGCAATATAATG
rs7140838	ACGTTGGATGCCCTCGTTTCTGAAGAATACC	ACGTTGGATGCTGACTGAACAGGTTATTGG
rs7141127	ACGTTGGATGCCCTGTTTCTGAAGAAATACC	ACGTTGGATGGAGACTGAACAGGTTATTGG
rs6574392	ACGTTGGATGAGAAAATAGCATAGGCTGGG	ACGTTGGATGAATGATCCATCCCTCCAG
rs8003691	ACGTTGGATGACTGAAGTCAAGTGAAGGCC	ACGTTGGATGTTAGGCCCTATACATGAGG
rs8003979	ACGTTGGATGCACAAAACCACTTCTGAAGC	ACGTTGGATGGGGCCCTAATTTTCCTTTGC
rs8010541	ACGTTGGATGCACTTTTCTTGCTAGCTTC	ACGTTGGATGCAGAAATGGCTAAAACCTGAAC
rs8016416	ACGTTGGATGTGCCCATAACTTCTTTGAC	ACGTTGGATGGCCACGGAACTCCTATATAGA
rs8016175	ACGTTGGATGTTTGAAGCACTGAGTGAAGTGA	ACGTTGGATGTCTTAACCGTGAGTGATCTG
rs7154571	ACGTTGGATGATGTGAGGAGCACCTCTGCC	ACGTTGGATGCTCTTCCCTTCTCAGACGG
rs7158826	ACGTTGGATGCACCTCCCTCCTGGACGGG	ACGTTGGATGGCCACCCCGCTGTGAGAGG
rs7159310	ACGTTGGATGACCCCGCTGTGAGAAGGGAGA	ACGTTGGATGCCATCTCCCTGTGACAGGG
rs7401900	ACGTTGGATGCCCAACAGCCTCATGTGAGAAC	ACGTTGGATGCTTTTCCCCACATTTCGCC
rs7160355	ACGTTGGATGTCACTTGTGTTTATCTGCTGAC	ACGTTGGATGTTATTGATCACTTCTGGGTG
rs2032781	ACGTTGGATGTATATCACTGTAGTAACAGC	ACGTTGGATGCCAAGATATATATACAAAG
rs6574394	ACGTTGGATGACCACAACCCAGCCTATTGTG	ACGTTGGATGTTATGTGCAAGCCCTGGGAG
rs8007598	ACGTTGGATGCTGGCAAAAGTCTTTAACAC	ACGTTGGATGTTGTCTTAAAGTCACGAATG
rs2267767	ACGTTGGATGGTTTACCATGTTAGCCAGG	ACGTTGGATGTAATCCAGCACTTTGGGAG
rs6574395	ACGTTGGATGAACCTTGAACCTTTGGGCTC	ACGTTGGATGAAAAAATACCCCGGGCATGG
rs7160356	ACGTTGGATGGAATACCTCTCTGCTTCTG	ACGTTGGATGAGATCAGGTGTAGATCAGG
rs7492334	ACGTTGGATGGCCTTTGCATTGGCTATTTG	ACGTTGGATGTAGAAAGCAGTCATGGGAG
rs4359361	ACGTTGGATGGTAGTATTGCTTAGTACAC	ACGTTGGATGTTCTAAGCCTGAATGTTTCC
rs4605089	ACGTTGGATGAATACCTATGAGATCTCAGG	ACGTTGGATGCTTGTAACTCTTTAACTAC
rs7146446	ACGTTGGATGATTCACCTTTTACAAGACCTC	ACGTTGGATGGCATATTGTACTTAGGAACCTC
rs4346144	ACGTTGGATGATTCACTTTTACAAGACCTC	ACGTTGGATGGCATATTGTACTTAGGAACCTC
rs7148078	ACGTTGGATGGTGTCAAGATTGAGGCTTG	ACGTTGGATGCCAAGAGAAATAAGCTGAGAG
rs7148286	ACGTTGGATGGTGGTCATTAAAGCTTGGCAG	ACGTTGGATGTGCTATGGATGCTGCTTGAG
rs3783980	ACGTTGGATGTTTTTGGCCAGGTAAGAC	ACGTTGGATGTGGTGCTTTTGTCTCTCTG
rs1549119	ACGTTGGATGTTTCACTTCTCTCTGCCCTCC	ACGTTGGATGGTGAAGGCCAGTCATATTGC
rs1984925	ACGTTGGATGAAGTAGCCAGGATTACAGGC	ACGTTGGATGCCAGCCTAGCAACATGGTG
rs1477261	ACGTTGGATGCAGGGTTATGTGGTATTATC	ACGTTGGATGGGGAAGTAAAGATAAGAG
rs8016447	ACGTTGGATGAATACAGACGTGTGCCACC	ACGTTGGATGTGACACAGAGAGACTCTGTC
rs7494044	ACGTTGGATGAATACAGACGTGTGCCACC	ACGTTGGATGTGACACAGAGACTCTGTC
rs2023288	ACGTTGGATGGAGAAAATGTGATTGATTG	ACGTTGGATGGCCATCAAAATGACTCTTAATC
rs7151685	ACGTTGGATGACAGTGCTGGCATTACTGGC	ACGTTGGATGTAAGATCGTCTGCCACTGC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2112135	ACGTTGGATGAGTGCAGTGGCCCAATCACA	ACGTTGGATGGTCTAGAGATCCGACTACTC
rs2161088	ACGTTGGATGTATAGGGTCTCACTCTTGCC	ACGTTGGATGAGGAGGATCACCTGAGCCTT
rs4903638	ACGTTGGATGATAGGGTGTCTACTGCGTTGG	ACGTTGGATGAGGCCCTAGGTGAGAAAGATTG
rs1477262	ACGTTGGATGATGCGGTGAGGAGAATGAAGG	ACGTTGGATGAAGGCTAGTGTTCAGGAAGG
rs1477263	ACGTTGGATGAACCTTCTGGAACACTAGCC	ACGTTGGATGCCTTGCTGCCCCATTTTAAG
rs1477264	ACGTTGGATGCGGTAGATAGAACCACCTCAG	ACGTTGGATGAAAGCGGGAGAGCACTTTAC
rs2277917	ACGTTGGATGGCATTGTTCGTAGCTGAAG	ACGTTGGATGTTGAACAGGAGTACCGTTTG
rs2277918	ACGTTGGATGTTACGTTCCCTACTCAGTCC	ACGTTGGATGACCTGTGCGTTTAAACGCC
rs2277919	ACGTTGGATGTTACGTTCCCTACTCAGTCC	ACGTTGGATGACCTGTGCGTTTAAACGCC
rs1978416	ACGTTGGATGAGGGCGTTTAAACGACACAGG	ACGTTGGATGCGGGTGAGAGGATATGGTTT
rs3759728	ACGTTGGATGATAGTCCCTCGCTGTTTGG	ACGTTGGATGAGAAAGCACTAGGCCCTTGG
rs6574399	ACGTTGGATGATGCTCTGATGCCATTATGC	ACGTTGGATGAGGGCAGCTAAACACATCC
rs1715536	ACGTTGGATGGAGGAAGACTCGGTCTAAAA	ACGTTGGATGAACAACCTGACACTAGGTGC
rs17156186	ACGTTGGATGATTACGGGTATGAGCCACTG	ACGTTGGATGGAACCTGACATTAAGGCTCG
rs7142390	ACGTTGGATGTAATCAAGACAGTGTGGTAC	ACGTTGGATGGCGTTTATTTCAGGACTCTC
rs7145875	ACGTTGGATGGTCCCTTGAAGCACAAAACC	ACGTTGGATGCTTCATGATCTTGGATTGGC
rs8014635	ACGTTGGATGGTCTTCTCACTCAAGAACAC	ACGTTGGATGCAACAGCAGCAAGCTCCAAC
rs8015938	ACGTTGGATGCCCTGAACCTCAAGTGATCTG	ACGTTGGATGATCTAAACAGGTGTCTGGC
rs8015313	ACGTTGGATGAAAAGTATTCTGTACCTGG	ACGTTGGATGGTCACTCATTTATAGGCAG
rs8006315	ACGTTGGATGTCACTTGAGGTCAGGAGTTC	ACGTTGGATGCCATTGCGCTGGCTAAGTTTG
rs6574400	ACGTTGGATGTTATCCCTTCTGCCAGTG	ACGTTGGATGCCCTTGAACCTTCTACCCAG
rs7140816	ACGTTGGATGCAATAGAGTGAAGCTCTGTC	ACGTTGGATGATTCATGAGCCTCTCTTTAC
rs4566078	ACGTTGGATGTAGAGTCTTGTCTGTCAAC	ACGTTGGATGAGGAGAATCGCTGAACCCCA
rs7141050	ACGTTGGATGATATGTTATACATATTAGTC	ACGTTGGATGCAAGTTGACCATATATCAACTC
rs3049356	ACGTTGGATGTACCACTGGCAGAGTAGAAG	ACGTTGGATGCACATGTTTGGGTACTGAG
rs4903639	ACGTTGGATGAGCGAGACTCCGCTCTCAAAA	ACGTTGGATGTCAAGGATAGCCTTGACTGG
rs4903641	ACGTTGGATGACTCCAACCTGGGCAACAGA	ACGTTGGATGCTGGCTCCAGCACACTATC
rs2364838	ACGTTGGATGTGTAGTCCAGCTACTTGTG	ACGTTGGATGTGATCATAGCTCACTGCAGC
rs2364839	ACGTTGGATGTGGGCAACATAGCAAGATCC	ACGTTGGATGCCTACAAGTAGCTGGGACTA
rs4632066	ACGTTGGATGGAGAAAAAGAGATGGAGGG	ACGTTGGATGGCCCTGACTGTGTTTTATG
rs2112136	ACGTTGGATGTTTCTTGGGGCACTAAGGCTC	ACGTTGGATGTACACGCCCTGAAGGAATTG
rs4641655	ACGTTGGATGCGATAGAGCAACCTGTCTC	ACGTTGGATGGCCCTACACAGGATCAAG
rs4635269	ACGTTGGATGAAAGTGCTGGGATTACAGGC	ACGTTGGATGCTTGACGACATATTTCTGAGG
rs4570764	ACGTTGGATGCTTGACGACATATTTCTGAGG	ACGTTGGATGAAAGTGCTGGGATTACAGGC
rs759808	ACGTTGGATGATGCTGTGATCATGCCAC	ACGTTGGATGCTGAACTGATTTGTGCTCC
rs7150531	ACGTTGGATGATGTGCGGTGTGAAGCAAAG	ACGTTGGATGTTGTTTGGCCTGGTCTGATG
rs7154968	ACGTTGGATGTACCCAGGTAACAAACCTGC	ACGTTGGATGTCCCTATAAGGGTTTCAGG
rs7146657	ACGTTGGATGTAGTACCTGGGACTACAGG	ACGTTGGATGTAACAGCGGTGAACACCTGC
rs7145859	ACGTTGGATGAGGCAAGGAGAAATGGCGTAA	ACGTTGGATGTTTGTGAGACGGAAGTCTG
rs4903643	ACGTTGGATGTATTCATGCTGTCTGCCTC	ACGTTGGATGAGTTGACCTTAAAGGCTGGG
rs717682	ACGTTGGATGTTTAAAGGCAGAGGCTGAGG	ACGTTGGATGAAGTGCAAGTGGCCTGATCTC
rs717683	ACGTTGGATGTTGGCAAAAAGGTTGAGGCG	ACGTTGGATGTGATGATGCCACAGGGAATG
rs1477259	ACGTTGGATGTGACTGAGACTACCTTCACC	ACGTTGGATGAAGTGCTCAGTAGGTTGAC
rs8019064	ACGTTGGATGCCCTTGACGCAAACTTCAGAG	ACGTTGGATGTGACTGAGACTACCTTCACC
rs8018971	ACGTTGGATGATGGTCTCACTCTGTCACTC	ACGTTGGATGAATTTGTTGAGCCCAAGGAG
rs1477260	ACGTTGGATGAGTGTCTATGTTAGCAAGGAC	ACGTTGGATGTGCCACTGTTTCCCATAGG
rs5809851	ACGTTGGATGACAGAGAGTGTTCAGCACAG	ACGTTGGATGTTGGGCAACAGAGAGAGACT

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs1985149	ACGTTGGATGACTGAATCTTTGCCTCCCG	ACGTTGGATGGTGGCACTTATGTAGTCC
rs1008988	ACGTTGGATGAGTGTGTCTCAGGGAATGTG	ACGTTGGATGCGCTGGCAATTGTCTCTGCG
rs1008989	ACGTTGGATGGGAATAGCAAGTGTAAACGGC	ACGTTGGATGACTCCAACCGCATCAGCTTC
rs8018222	ACGTTGGATGATCCTCCATATGCTGAACGC	ACGTTGGATGAAGGTGGAACGAGAGACTTG
rs1006040	ACGTTGGATGTTTACGCTCTCTCTCTGTTC	ACGTTGGATGCTTGGAGCCAGGAGTTCAA
rs1006039	ACGTTGGATGTGAAGCTGGGAGTTAGAGAC	ACGTTGGATGCCACCATGCCAGCTAATT
rs1006038	ACGTTGGATGATAAGCCACTGTGCTCAGTC	ACGTTGGATGGGTAGGGTTTATTAAGTGCC
rs8009784	ACGTTGGATGTGTTTGGCTATGCTTTGCC	ACGTTGGATGTGACAGAGCGAGACTTTGTC
rs4903644	ACGTTGGATGTTGCAGTGAGCTGAGATTGG	ACGTTGGATGGTGAATGAATGAATAAGGGCC
rs7149496	ACGTTGGATGACAACACACAGTACTGGACC	ACGTTGGATGTGGGTGCATGTTAGAAACGC
rs6574402	ACGTTGGATGCAGGTCTTTGTCTGACAAG	ACGTTGGATGGGATGTGCGATTTGATCTG

TABLE 22

dbSNP rs#	Extend Primer	Term Mix
rs7143926	ACCCAAAATTAAGGCCAAAATGG	CGT
rs1549071	CACACACATATATACACACACA	ACG
rs8012858	CACACATATATACACACACACA	ACG
rs7155611	GGCACCGTTCTCTCTCTCA	ACT
rs176941	CTGGGCCCTCAGTTTACTCAT	CGT
rs176942	AATAGGTTGGTTTGTGCCCC	ACT
rs176943	CCCGTAGTCCCTGTGAAAC	ACT
rs176944	AAAAGTCCACTAATCCTTCCAA	CGT
rs4365221	GAGGGCAACTCAACACATTTTA	ACG
rs3168952	TTGGTGGTGAGATGGACAGA	ACT
rs176945	TACTATACACTCACAAAATTGTT	ACT
rs176946	TTGTATAACAAAATACCACAAGC	ACT
rs176947	GGCGCCCGCCACTACGC	ACG
rs176948	AAAACAGACCTCAGTCTACA	ACT
rs176949	CTCCCTGCAGTTCCTTGTTA	CGT
rs176950	GGAGTGACATCCCATTACTTT	ACG
rs176951	TCCTCCCTCCTTGGGTG	ACT
rs7156905	CTGTCTCAAAAAAGGAACCAG	ACT
rs3217197	CTCCAGCCTGAGTGAGAGA	ACT
rs2270443	CAGGTGTGGTGGCTCATGC	ACG
rs176952	CAGTGAATCGCTGCCTCC	ACG
rs176953	GGACCAGCCTGGCCAACAT	ACT
rs176954	GCAGTGGCTCAATCCCAGC	CGT
rs176955	CTGCCCTCCAGCCCTTC	ACT
rs3214416	CATCACGTGTTCTAATGAAAA	CGT
rs176956	AAGCCACCATGCCAGCC	ACT
rs2544566	CATCTGGGCCTCCCAAGTA	ACT
rs2544567	CATCTGGGCCTCCCAAGT	ACT

dbSNP rs#	Extend Primer	Term Mix
rs176957	GGTCTCGATCTCCTGACCT	ACG
rs176958	GGAGTTTGTCTCTTGTGGCC	ACT
rs176959	TTTTATTGTCCCTTGTCTTTC	ACT
rs1802227	AATAGTTGCACCAAGCAAGAG	ACT
rs176961	TATGGCAAAACCTGTCTACA	ACT
rs176962	GGCTCAGCCTGTAACTCTA	ACT
rs7401285	GGGACTCCATATCAGAAAACA	CGT
rs176963	GAAAACACACGCGGGCGC	ACT
rs176964	CAGTGTCTACAAAAGTGCCT	ACG
rs4903631	CTTGAGACAAGATGAAACAGTT	ACG
rs4903632	ATCCTCAGGGAAACGAAAATTA	CGT
rs176965	ATAAGGGCTGCCAGCTTGAT	ACT
rs4903633	TAGCAATTTTATATCTCAGCATG	ACT
rs176966	ACCACACCCAGCTAATTTTG	ACG
rs176968	TCACACCTGTGACTCCAGC	CGT
rs176969	CCGATTTACTGCAATTGCAATTC	CGT
rs176970	GTACAGTGGGTGAATAGTTA	ACT
rs7149198	GATATTACTCAGCCATAAAAAAG	ACT
rs7147918	GGATATTACTCAGCCATAAAAAA	ACT
rs7148685	TTGAGACCGTCTATTGATC	ACT
rs1184232	GAATGGAAGAAAATGGTTGCAAA	CGT
rs1184233	TGCCCAGCCTCTTCAATTAC	ACT
rs1184234	TACCAGCACTTTGGGAGGC	ACG
rs7401988	CCACGCTGGCTAATTTTITTT	CGT
rs176974	CTGGGCAACAAAGCAAGACT	ACG
rs6574390	AAATTAGCTGGGTATGATGGC	ACT
rs176975	AATCTAGGTGGTAGGAGATAC	ACT
rs176976	TATAATCTTTTCTGCTTCTGTA	ACT
rs176977	TCAGCCTGGGCAACAAGAG	ACG
rs8013727	CCTAACCATAGAAAGATAATTAGAA	ACT
rs176978	ATGTTTTGTACCTCTTGTAC	ACG
rs2111829	GAATTTTGCTTGGTGAACAAAAT	ACT
rs176980	GCTATGAACGCCATTTATGTA	ACT
rs5809848	TGGGTTCTGAAATCCTGCTG	CGT
rs5809849	CGTGTCAAGTTCTTTTTTTTTT	ACT
rs4383070	GCCTCCTGAGTAGCTGGG	CGT
rs7493652	TATACCACAAGTAACTGTTAATTT	ACG
rs2112133	CCACAACCTGGCTAATTTTTTGT	CGT
rs1963833	CTGGGTGACAGAGCAAGAC	CGT
rs6574391	TGGAGAAGTGATAAACTC	ACG
rs7155062	ATAACCCCTCAAATGAGCATCA	ACT
rs4899674	GGCAAAATGGCTGGGGAG	ACG
rs8022516	AGGCATTCTTTTGGGTATAGTA	ACG
rs7140838	TCTGAAGAATACCAGACCTCT	CGT

dbSNP rs#	Extend Primer	Term Mix
rs7141127	CTGAAGAATACCAGACCTCTC	ACT
rs6574392	CTGGGCACAGCGACTCAC	ACT
rs8003691	TGAAGGCCCTCCATGGTATAG	ACT
rs8003979	TCTGAAGCCAGTGAGGAAGT	ACT
rs8010541	GCTAGCTTCAACTCTCCTGAT	ACT
rs8016416	TAACCTCCTTTGACTTGTCTTTT	CGT
rs8016175	GTCTGCAATCCCGGCACCT	ACG
rs7154571	GTGAGGAGCGTCTCTGCC	ACG
rs7158826	TCGCTCCTCACTTCCGAGA	ACG
rs7159310	CATCTGGGAAGTGAGGAGC	ACT
rs7401900	TGAGAACAGCCATGATGAC	ACT
rs7160355	CCTGCCAAATCCCCCTCTC	ACG
rs2032781	GAGAAAAGCGGGCAGGACT	ACT
rs6574394	CACCCAGCCTATTGTATAATT	ACT
rs8007598	CTCTTAACACATTTTTTACAGCA	ACG
rs2267767	CTGACCTCGTGATCTGCC	ACT
rs6574395	GGCTCAGGCGATCATCGTA	ACG
rs7150066	CTGCCACCCAAAGTGCTGG	ACT
rs7492334	TTGTGTGTGTGTGTGTGTGG	ACT
rs4359361	GCTTAGTACATTTAAACATGAT	ACT
rs4605089	TCAGGAACACCGCTTAATTTTT	ACG
rs7146446	CAAGACCTCTTTAAGTAATACTC	ACG
rs4346144	AGACCTCTTTAAGTAATACTCC	ACT
rs7148078	GGCTTGGGTACGGGAAGC	CGT
rs7148286	CATTAAGCTTGCCAGAAAATCA	ACG
rs3783980	CATCTTCCTCTGCCTCCCA	ACG
rs1549119	CTTCCTCTGCCTCCCATAAAT	CGT
rs1984925	CAGGCACGTGCCACCACA	ACG
rs1477261	AGGAGGAGCCCCAAATATGAAA	CGT
rs8016447	CACACCTGGCCATGCTTCC	ACT
rs7494044	CCTGGCCATGCTTCCGTATT	ACG
rs2023288	AATTGTGATTGATTGATTGCGAT	CGT
rs7151685	GTGAGCCACCATCATCTG	ACT
rs2112135	TCAGGTGATCCTCCTGCCT	ACG
rs2161088	CCAATCACAGCCCATGCA	ACT
rs4903638	GCCAGAGTGGTCTCCAAC	ACT
rs1477262	AGAGCTCAAGCTGATGTCCT	ACT
rs1477263	TTTTCTGTGAGTTCCGATG	ACT
rs1477264	ACCACCTCAGTTTTGCTGTTT	ACG
rs2277917	CCTTGATAACCGCTTGGTCT	ACT
rs2277918	AAAAGCTTCCCGGGGACAG	CGT
rs2277919	AGCTTCCCGGGGACAGCT	ACT
rs1978416	TGAGACTAGCTAATGGAGAGT	ACG
rs3759728	AGCAAATCTACTGCAAACGTG	CGT

dbSNP rs#	Extend Primer	Term Mix
rs6574399	AAGTAGAGCTGCTCCACC	CGT
rs7155336	GAAGACTCGGTCTAAAAA	ACT
rs7156186	GCCACTGCACCTGGCCG	ACT
rs7142390	TGGTACTGGCATAAGGATAGA	ACG
rs7145875	CACAAAACCTTAACCTTTTGATT	ACT
rs8014635	CAAGAACTGGTTTGGTTT	ACT
rs8015938	CTCAAGTGATCTGCCTGCC	ACT
rs8015313	GATTCTGTACCTGGTTGATCAT	ACT
rs8006315	AACATGGTGAAGCCCCATCT	CGT
rs6574400	GAGATCGCCAGAGACACCA	ACG
rs7140816	TGAGACTCTGTCTCAAATACTA	CGT
rs4566078	CTCAGCTCACTGCAACCTC	ACG
rs7141050	AGCACATAGTAAGTGCCCTAT	ACT
rs3049356	GAATAGTGAAGGTATTGAAATA	ACT
rs4903639	GAGACTCCGTCTCAAAAAA	CGT
rs4903641	GGCAACAGAGCGGAGACTCC	ACT
rs2364838	CCAGCTACTTTGTGAGGCCAA	ACT
rs2364839	AGCCAGACGTGGTGGCAC	ACT
rs4632066	GAGATGGAGGGGGAGCCT	ACT
rs2112136	GGGACTAAGGCTCGCATCC	ACT
rs4641655	GGATTTCTGGGTCCCACTC	ACG
rs4635269	AGCCACCGCGCCCGGCC	ACT
rs4570764	GTGATTATTGGCCGGGCGC	ACT
rs759808	TGCACCACACAGCTGGG	CGT
rs7150531	AGCAAAGTTAATGGGAGGCC	ACT
rs7154968	AACAAACCTGCATATGTACCC	ACT
rs7146657	CACCCACACCCCGCCC	ACT
rs7145859	CGGGAGGTGGAGCTTGCA	ACT
rs4903643	GCTCCCTTCTGTCTACTGC	ACT
rs717682	AGGCTGAGGCAGGAGAATC	ACT
rs717683	AAAAGTGGAGGCCAAAGAC	ACT
rs1477259	CGGAATAATTATATCGCTCT	ACT
rs8019064	CAGAGGCAGATATAATTATCC	ACT
rs8018971	GCTACTCAGGCTGGAGTGC	CGT
rs1477260	GACGAGGAGGAAAGCCATC	ACT
rs5809851	CACAGCAGTGTCTTTTTTTT	ACT
rs1985149	TTCTTCTCCCTCAGCCTCC	ACG
rs1008988	GGGGATGACCTCTCTGGAG	ACT
rs1008989	GCCAGCTTGGCAGATTGAG	CGT
rs8018222	TGCTGAACGCTGGTCCCC	CGT
rs1006040	TGGAGTGCACTGGCAAGAC	ACG
rs1006039	CATAGCCAGACCTATGAGA	ACG
rs1006038	ACTGTGCTCAGTCTATGCTG	ACG
rs8009784	ATGCTTTGCCTAAAGTGGTG	ACT

dbSNP rs#	Extend Primer	Term Mix
rs4903644	GCCTGGGCAACAGAGCAAG	ACT
rs7149496	GATTCTGTAAGTCTGGTATGAG	ACT
rs6574402	CTGACAAGAAAATGACTGCATA	ACT

Genetic Analysis

[0242] Allelotyping results are shown for cases and controls in Table 23. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where "AF" is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs7143926 has the following case and control allele frequencies: case A1 (A) = 0.75; case A2 (T) = 0.25; control A1 (A) = 0.71; and control A2 (T) = 0.29, where the nucleotide is provided in paranthesis. Some SNPs are labeled "untyped" because of failed assays.

TABLE 23

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs7143926	218	76161268	A/T	0.25	0.29	0.216
rs1549071	1440	76162490	C/T	0.15	0.20	0.098
rs8012858	1442	76162492	C/T	0.93	0.95	0.335
rs7155611	2611	76163661	C/T	0.02	0.02	0.949
rs176941	4317	76165367	A/C	0.31	0.35	0.271
rs176942	4724	76165774	A/G	0.02	0.02	0.911
rs176943	4788	76165838	G/T	0.13	0.18	0.037
rs176944	5202	76166252	G/T	0.09	0.14	0.107
rs4365221	5780	76166830	C/T			
rs3168952	5974	76167024	C/T			
rs176945	6644	76167694	C/G	0.95	0.96	0.801
rs176946	7430	76168480	A/G	0.10	0.15	0.054
rs176947	7938	76168988	C/T	0.10	0.08	0.473
rs176948	8095	76169145	C/T	0.31	0.35	0.132
rs176949	8183	76169233	A/C	0.03	0.02	0.887
rs176950	8312	76169362	C/T	0.78	0.70	0.008
rs176951	8352	76169402	A/C			
rs1756905	9348	76170398	C/T	0.89	0.90	0.794
rs3217197	9378	76170428	-T/CTC	0.29	0.35	0.036
rs2270443	9617	76170667	A/G	0.39	0.34	0.176
rs176952	9727	76170777	C/T	0.17	0.24	0.018
rs176953	9834	76170884	C/T			
rs176954	9899	76170949	G/T	0.43	0.52	0.010
rs176955	10211	76171261	C/T	0.12	0.18	0.028
rs3214416	10377	76171427	-T	0.91	0.89	0.544
rs176956	10695	76171745	C/T	0.51	0.49	0.492
rs2544566	10729	76171779	C/G			
rs2544567	10730	76171780	C/T			
rs176957	11433	76172483	A/G			
rs176958	11951	76173001	C/G	0.02	NA	NA



dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs176959	12697	76173747	C/T	0.30	0.34	0.147
rs1802227	12982	76174032	A/C	0.92	0.95	0.332
rs176961	14419	76175469	C/T	0.51	0.47	0.158
rs176962	14501	76175551	C/T	0.82	0.79	0.192
rs176963	14983	76176033	A/C			
rs176963	15280	76176330	C/T	0.51	0.46	0.155
rs176964	15475	76176525	A/G	0.53	0.49	0.197
rs4903631	15888	76176938	A/G			
rs4903632	15976	76177026	A/T			
rs176965	16307	76177357	A/C	0.55	0.52	0.368
rs4903633	16442	76177492	A/C	0.83	0.83	0.970
rs176966	17255	76178305	C/T			
rs176968	18948	76179998	G/T	0.23	0.27	0.246
rs176969	19435	76180485	A/T	0.14	0.20	0.052
rs176970	19753	76180803	C/T	0.35	0.38	0.328
rs7149198	20021	76181071	C/T			
rs7147918	20022	76181072	A/C			
rs7148685	20503	76181553	A/G	0.19	0.18	0.669
rs1184232	20590	76181640	G/T	0.16	0.19	0.316
rs1184233	21804	76182854	G/T	0.36	0.36	0.895
rs1184234	21919	76182969	C/T	0.36	0.35	0.797
rs7401998	21990	76183040	A/T			
rs176974	22412	76183462	A/G			
rs6574390	22536	76183586	C/T			
rs176975	23432	76184482	A/G	0.18	0.23	0.147
rs176976	23468	76184518	G/T	0.86	0.80	0.087
rs176977	23772	76184822	C/T	0.42	0.41	0.794
rs8013727	24325	76185375	C/T			
rs176978	24773	76185823	C/T	0.10	0.12	0.512
rs2111829	26274	76187324	C/T	0.02	NA	
rs176980	27440	76188490	C/G	0.79	0.73	0.018
rs5809848	28561	76189611	-ACAG	0.11	0.16	0.091
rs5809849	30071	76191121	-A	0.60	0.57	0.355
rs4383070	31764	76192814	A/T			
rs7493652	33008	76194058	C/T			
rs2112133	35310	76196360	A/T			
rs1963833	35460	76196510	A/C			
rs6574391	37112	76198162	A/G	0.69	0.63	0.064
rs7155062	37285	76198335	A/G	0.17	0.18	0.878
rs4899674	37747	76198797	C/T	0.57	0.52	0.201
rs8022516	38057	76199107	C/T	0.57	0.51	0.135
rs7140838	38859	76199909	A/C	0.17	0.17	0.957
rs7141127	38860	76199910	A/G			
rs6574392	39525	76200575	A/G	0.27	0.32	0.099
rs8003691	40216	76201266	A/G	0.70	0.63	0.029
rs8003979	40281	76201331	C/T	0.10	0.15	0.024
rs8010541	41453	76202503	C/G	0.38	0.38	0.993
rs8016416	42091	76203141	A/T	0.09	0.14	0.035
rs8016175	42513	76203563	A/G			
rs7154571	42935	76203985	C/T			
rs7158826	42985	76204035	A/G			
rs7159310	43003	76204053	A/G	0.62	NA	
rs7401900	43281	76204331	A/G			
rs7160355	43716	76204766	C/T			
rs2032781	43866	76204916	A/G	0.80	0.74	0.047
rs6574394	44234	76205284	G/T	0.61	0.54	0.091
rs8007598	44596	76205646	A/G	0.09	0.10	0.734

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2267767	44871	76205921	C/T			
rs6574395	45005	76206055	A/G	0.10	0.14	0.203
rs7150066	45282	76206332	A/C	0.91	NA	
rs7492334	47178	76208228	A/C			
rs4359361	47816	76208866	G/T			
rs4605089	47887	76208937	A/G			
rs7146446	48134	76209184	C/T	0.09	0.09	0.981
rs4346144	48135	76209185	A/G	0.83	0.85	0.368
rs7148078	48276	76209326	G/T	0.44	0.50	0.098
rs7148286	48400	76209450	C/T	0.96	0.96	0.893
rs3783980	48798	76209848	A/G	0.15	0.20	0.073
rs1549119	48803	76209853	A/T	0.18	0.25	0.027
rs1984925	49146	76210196	C/T	0.04	0.04	0.882
rs1477261	49969	76211019	A/T			
rs8016447	51059	76212109	A/G	0.10	0.15	0.049
rs7494044	51064	76212114	C/T			
rs2023288	53285	76214335	A/T	0.97	0.98	0.774
rs7151685	54560	76215610	C/T			
rs2112135	54748	76215798	A/G	0.05	NA	
rs2161088	54785	76215835	C/G			
rs4903638	55102	76216152	C/G	0.59	0.59	0.975
rs1477262	55644	76216694	A/G	0.12	0.17	0.040
rs1477263	55705	76216755	G/T	0.18	0.23	0.057
rs1477264	55841	76216891	A/G	0.45	0.42	0.271
rs2277917	56623	76217673	C/G	0.30	0.36	0.039
rs2277918	56825	76217875	A/C	0.49	0.45	0.232
rs2277919	56827	76217877	A/G	0.20	0.17	0.310
rs1978416	56892	76217942	C/T	0.79	0.73	0.074
rs3759728	59150	76220200	A/T	0.13	0.18	0.083
rs6574399	59958	76221008	A/T	0.33	0.36	0.396
rs7155336	60231	76221281	C/T	0.25	0.28	0.250
rs7156186	60524	76221574	A/G	0.85	0.85	0.965
rs7142390	61871	76222921	C/T			
rs7145875	62226	76223276	C/T			
rs8014635	63230	76224280	G/T	0.07	0.11	0.062
rs8015938	63468	76224518	G/T	0.08	0.07	0.693
rs8015313	63787	76224837	C/T	0.67	0.71	0.135
rs8006315	65732	76226782	A/C			
rs6574400	65989	76227039	A/G	0.75	0.70	0.099
rs7140816	68832	76229882	G/T	0.54	0.48	0.095
rs4566078	69904	76230954	C/T			
rs7141050	70365	76231415	A/G			
rs3049356	70886	76231936	-TATC	0.64	0.69	0.091
rs4903639	73088	76234138	A/T			
rs4903641	73103	76234153	C/T	0.54	0.66	~0.0001
rs2364838	75934	76236984	C/T			
rs2364839	75966	76237016	C/T	0.18	0.18	0.988
rs4632066	76273	76237323	C/T	0.66	0.66	0.961
rs2112136	77943	76238993	C/T	0.70	0.64	0.064
rs4641655	78466	76239516	C/T	0.52	0.48	0.174
rs4635269	78861	76239911	C/T			
rs4570764	78872	76239922	A/G	0.55	0.68	~0.0001
rs759808	79836	76240886	G/T	0.12	0.18	0.043
rs7150531	80908	76241958	C/T	0.33	0.31	0.491
rs7154968	81509	76242559	C/G	0.03	NA	
rs7146657	83576	76244626	C/T	0.57	NA	NA
rs7145859	83662	76244712	C/G			

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs4903643	83782	76244832	C/T	0.10	0.14	0.074
rs717682	84282	76245332	G/T	0.11	0.13	0.624
rs717683	84444	76245494	A/G	0.79	0.75	0.121
rs1477259	85129	76246179	C/G	0.11	0.16	0.022
rs8019064	85151	76246201	A/G	0.90	0.93	0.192
rs8018971	85296	76246346	A/C			
rs1477260	85809	76246859	C/G	0.12	0.16	0.085
rs5809851	86387	76247437	-/T	0.30	0.30	0.993
rs1985149	86494	76247544	A/G	0.22	0.23	0.892
rs1008988	89786	76250836	A/G	0.61	0.58	0.380
rs1008989	89894	76250944	A/T	0.14	0.18	0.172
rs8018222	90122	76251172	G/T			
rs1006040	92067	76253117	A/G	0.13	0.18	0.092
rs1006039	92187	76253237	C/T	0.06	0.10	0.133
rs1006038	92312	76253362	A/G	0.19	0.24	0.114
rs8009784	92824	76253874	G/T	0.13	0.18	0.037
rs4903644	93733	76254783	C/T	0.41	0.38	0.383
rs7149496	96553	76257603	C/G			
rs6574402	96941	76257991	A/C	0.12	0.17	0.037

[0243] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1C for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1C can be determined by consulting Table 23. For example, the left-most X on the left graph is at position 76161268. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0244] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise

level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than  $10^{-8}$  were truncated at that value.

[0245] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

#### Example 7

##### *In Vitro* Production of Target Polypeptides

[0246] cDNA is cloned into a pIVEX 2.3-MCS vector (Roche Biochem) using a directional cloning method. A cDNA insert is prepared using PCR with forward and reverse primers having 5' restriction site tags (in frame) and 5-6 additional nucleotides in addition to 3' gene-specific portions, the latter of which is typically about twenty to about twenty-five base pairs in length. A Sal I restriction site is introduced by the forward primer and a Sma I restriction site is introduced by the reverse primer. The ends of PCR products are cut with the corresponding restriction enzymes (*i.e.*, Sal I and Sma I) and the products are gel-purified. The pIVEX 2.3-MCS vector is linearized using the same restriction enzymes, and the fragment with the correct sized fragment is isolated by gel-purification. Purified PCR product is ligated into the linearized pIVEX 2.3-MCS vector and *E. coli* cells transformed for plasmid amplification. The newly constructed expression vector is verified by restriction mapping and used for protein production.

[0247] *E. coli* lysate is reconstituted with 0.25 ml of Reconstitution Buffer, the Reaction Mix is reconstituted with 0.8 ml of Reconstitution Buffer; the Feeding Mix is reconstituted with 10.5 ml of Reconstitution Buffer; and the Energy Mix is reconstituted with 0.6 ml of Reconstitution Buffer. 0.5 ml of the Energy Mix was added to the Feeding Mix to obtain the Feeding Solution. 0.75 ml of Reaction Mix, 50  $\mu$ l of Energy Mix, and 10  $\mu$ g of the template DNA is added to the *E. coli* lysate.

[0248] Using the reaction device (Roche Biochem), 1 ml of the Reaction Solution is loaded into the reaction compartment. The reaction device is turned upside-down and 10 ml of the Feeding Solution is loaded into the feeding compartment. All lids are closed and the reaction device is loaded into the RTS500 instrument. The instrument is run at 30°C for 24 hours with a stir bar speed of 150 rpm. The pIVEX 2.3 MCS vector includes a nucleotide sequence that encodes six consecutive histidine amino acids on the C-terminal end of the target polypeptide for the purpose of protein purification. Target polypeptide is purified by contacting the contents of reaction device with resin modified with  $\text{Ni}^{2+}$  ions. Target polypeptide is eluted from the resin with a solution containing free  $\text{Ni}^{2+}$  ions.

### Example 8

#### Cellular Production of Target Polypeptides

[0249] Nucleic acids are cloned into DNA plasmids having phage recombination sites and target polypeptides are expressed therefrom in a variety of host cells. Alpha phage genomic DNA contains short sequences known as attP sites, and *E. coli* genomic DNA contains unique, short sequences known as attB sites. These regions share homology, allowing for integration of phage DNA into *E. coli* via directional, site-specific recombination using the phage protein Int and the *E. coli* protein IHF. Integration produces two new att sites, L and R, which flank the inserted prophage DNA. Phage excision from *E. coli* genomic DNA can also be accomplished using these two proteins with the addition of a second phage protein, Xis. DNA vectors have been produced where the integration/excision process is modified to allow for the directional integration or excision of a target DNA fragment into a backbone vector in a rapid *in vitro* reaction (Gateway™ Technology (Invitrogen, Inc.)).

[0250] A first step is to transfer the nucleic acid insert into a shuttle vector that contains attL sites surrounding the negative selection gene, ccdB (e.g. pENTER vector, Invitrogen, Inc.). This transfer process is accomplished by digesting the nucleic acid from a DNA vector used for sequencing, and to ligate it into the multicloning site of the shuttle vector, which will place it between the two attL sites while removing the negative selection gene ccdB. A second method is to amplify the nucleic acid by the polymerase chain reaction (PCR) with primers containing attB sites. The amplified fragment then is integrated into the shuttle vector using Int and IHF. A third method is to utilize a topoisomerase-mediated process, in which the nucleic acid is amplified via PCR using gene-specific primers with the 5' upstream primer containing an additional CACC sequence (e.g., TOPO® expression kit (Invitrogen, Inc.)). In conjunction with Topoisomerase I, the PCR amplified fragment can be cloned into the shuttle vector via the attL sites in the correct orientation.

[0251] Once the nucleic acid is transferred into the shuttle vector, it can be cloned into an expression vector having attR sites. Several vectors containing attR sites for expression of target polypeptide as a native polypeptide, N-fusion polypeptide, and C-fusion polypeptides are commercially available (e.g., pDEST (Invitrogen, Inc.)), and any vector can be converted into an expression vector for receiving a nucleic acid from the shuttle vector by introducing an insert having an attR site flanked by an antibiotic resistant gene for selection using the standard methods described above. Transfer of the nucleic acid from the shuttle vector is accomplished by directional recombination using Int, IHF, and Xis (LR clone). Then the desired sequence can be transferred to an expression vector by carrying out a one hour incubation at room temperature with Int, IHF, and Xis, a ten minute incubation at 37°C with proteinase K, transforming bacteria and allowing expression for one hour, and then plating on selective media. Generally, 90% cloning efficiency is achieved by this method. Examples of expression vectors

are pDEST 14 bacterial expression vector with att7 promoter, pDEST 15 bacterial expression vector with a T7 promoter and a N-terminal GST tag, pDEST 17 bacterial vector with a T7 promoter and a N-terminal polyhistidine affinity tag, and pDEST 12.2 mammalian expression vector with a CMV promoter and neo resistance gene. These expression vectors or others like them are transformed or transfected into cells for expression of the target polypeptide or polypeptide variants. These expression vectors are often transfected, for example, into murine-transformed adipocyte cell line 3T3-L1, (ATCC), human embryonic kidney cell line 293, and rat cardiomyocyte cell line H9C2.

[0252] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the claims which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0253] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

#### Nucleotide and Amino Acid Sequence Examples

[0254] Table A includes information pertaining to the incident polymorphic variant associated with osteoarthritis identified herein. Public information pertaining to the polymorphism and the genomic sequence that includes the polymorphism are indicated. The genomic sequences identified in Table A may be accessed at the http address [www.ncbi.nih.gov/entrez/query.fcgi](http://www.ncbi.nih.gov/entrez/query.fcgi), for example, by using the publicly available SNP reference number (*e.g.*, rs552). The chromosome position refers to the position of the SNP within NCBI's Genome Build 34, which may be accessed at the following http address: [www.ncbi.nlm.nih.gov/mapview/map\\_search.cgi?chr=hum\\_chr.inf&query=](http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?chr=hum_chr.inf&query=). The "Contig Position" provided in Table A corresponds to a nucleotide position set forth in the contig sequence (see "Contig Accession No."), and designates the polymorphic site corresponding to the SNP reference number. The sequence containing the polymorphisms also may be referenced by the "Nucleotide Accession No." set forth in Table A. The "Sequence Identification" corresponds to cDNA sequence that encodes associated target polypeptides (*e.g.*, Q96FX2). The position of the SNP within the cDNA sequence is provided in the "Sequence Position" column of Table A. If the SNP falls within an exon, the corresponding amino acid position (and amino acid change, if applicable) is provided as well. Also, the allelic variation at the

polymorphic site and the allelic variant identified as associated with osteoarthritis is specified in Table A. All nucleotide and polypeptide sequences referenced and accessed by the parameters set forth in Table A are incorporated herein by reference. Genomic nucleotide sequences for *KIAA0296*, *chrom 4* and *SNW1* regions are set forth in SEQ ID NO: 1-3, respectively. A polymorphism in Table A designated by "AA" is present in the genomic nucleotide sequence of SEQ ID NO: 4, which follows Table A.

Table A

RS_ID	Chromosome	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
552	3	1627693	HS2_22673_34:16	16241963	XM_209584	Intergenic		Q96FX2	253631; 92106	R	[A/G]	A
12904	1	152323469	HS1_79549_34:1	1556529	NM_004428 NM_018845	mmu-utr locus-region		EFNA1 LOC5597	1942 55974	F	[A/G]	A
2282146	20	49881706	HS20_11519_34:8	14249192	NM_002827	coding-synon	303	PTPN1	5770	R	[C/T]	T
734784	20	44409056	HS20_11519_34:8	8776542	NM_002251	coding-nonsynon	V489I	KCN51	3787	R	[G/A]	G
1042164	19	13125398	HS19_11452_34:1	4527200	NM_004907	coding-nonsynon	V133A	ETR101	9592	R	[T/C]	T
749670	16	31124685	HS16_24968_34:1	2488292	NM_014699	coding-nonsynon	G327E	KIAA029	9726	F	[C/T]	T
955592	2	85570358	HS2_22340_34:13	64411756	NM_032213	coding-nonsynon	I70T	FLJ21977	84173	R	[T/C]	C
1143016	1	2209625	HS1_4507_34:16	317874	NM_002617	coding-synon	93	PEX10	5192	R	[T/C]	T
755248	1	39403404	HS1_33153_34:6	1601557	NM_024732	mmu-utr		FLJ14351	79787	R	[G/A]	G
1055055	1	42293810	HS1_33153_34:6	4486963	NM_173642	coding-synon	308	MGC478	284716	F	[A/G]	A
835409	1	59157267	HS1_33153_34:6	21350420	NM_018291	mmu-utr		FLJ10986	55277	F	[T/G]	G
927663	1	155318902	HS1_79549_34:1	4551942	AL138899	10739		LOC2540	254079	F	[T/G]	G
8162	2	42552684	HS2_22340_34:13	21394082	NM_004718	mmu-utr		COX7A2	9167	F	[A/G]	G
831038	2	170404402	HS2_5560_34:14	20411275	NM_004525	intron		LRP2	4036	F	[C/T]	T
33079	2	231577157	HS2_5560_34:14	81594030	NM_003113	intron		SP100	6672	R	[G/A]	A
1710880	3	11027339	HS3_22673_34:16	10982336	AC024910.5	71065		SLC6A1	6529	F	[C/A]	A
4472029	3	102614466	HS3_5769_34:14	7788711	NM_020357	intron		PONP	57092	F	[T/A]	A
799570	3	113007639	HS3_5769_34:14	18181884	NM_145753	intron		LL5beta	90102	F	[A/G]	G
1282730	3	113026933	HS3_5769_34:14	18201178	NM_018394	intron		FLJ11342	55347	R	[G/A]	A
1518875	3	188430421	HS3_5769_34:14	93604666	none	Intergenic		none	none	R	[T/C]	C
1568694	4	8584835	HS4_8464_34:15	505708	NM_003501	intron		ACOX3	8310	R	[G/A]	A



RS_ID	Chromosome	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
905042	4	16229175	Hs4_6473_34:15	7153650	none	intergenic		none	none	R	[A/T]	A
1957723	4	36919171	Hs4_16453_34:15	4170634	none	intergenic		none	none	R	[G/A]	A
794018	4	40066604	Hs4_16453_34:15	7318067	none	intergenic		none	none	R	[G/A]	G
707723	4	47757791	Hs4_6395_34:10	7544215	NM_006587	UTR/intergenic		PRSC	10899	F	[C/T]	T
893861	4	47872633	Hs4_6395_34:10	7659057	NM_000087	mRNA-utr		CNGA1	1259	R	[G/A]	G
1914903	4	125442695	Hs4_16510_34:16	49479005	none	locus-region		LOC377201	377201	R	[G/A]	A
2062232	4	163485648	Hs4_16762_34:16	24456077	NM_020116	intron		DKFZp566D234	56884	F	[C/T]	T
26609	5	53447781	Hs5_6588_34:13	3996639	NM_019087	intron		FLJ20051	54622	R	[A/T]	T
1370987	5	79365937	Hs5_6870_34:13	8673996	XM_293971	locus-region		LOC345778	345778	F	[A/G]	A
1012414	5	150535958	Hs5_29448_34:10	11670384	NM_001155	intron		ANXA6	309	R	[G/A]	G
435903	5	156713884	Hs5_23289_34:11	1542547	NM_014376	intron		CYFIP2	26999	R	[G/A]	G
1248	5	156742512	Hs5_23289_34:11	1571175	NM_014376	intron		CYFIP2	26999	F	[T/A]	T
703508	5	169791195	Hs5_23289_34:11	14625658	NM_004137	mRNA-utr		KCNMB1	3779	R	[G/A]	A
226485	6	1755892	Hs6_35042_34:3	1750892	NM_001500	intron		GNDS	2762	R	[G/C]	C
241448	6	32785099	NT_033951	4177541	NM_000544	UTR		TAP2	6891	F	[C/T]	C
763155	6	35240628	Hs6_7749_34:13	25909000	NM_152753	intron		CEGF3	222663	R	[A/C]	A
1040461	6	57102190	Hs6_7749_34:13	47852462	NM_016277	coding-nonsynon	S207G	RAB23	51715	R	[T/C]	C
462832	6	111731861	Hs6_25897_34:13	15855720	NM_002912	intron		REV3L	5980	R	[A/T]	A
804194	6	112053555	Hs6_25897_34:13	16177414	NM_002037	intron		FYN	2534	R	[T/C]	C
1022646	6	112074901	Hs6_25897_34:13	16198760	NM_002037	intron		FYN	2534	F	[A/G]	A
1569112	7	699341	Hs7_7976_34:14	250677	NM_006869	intron		CENTA1	11033	R	[G/A]	G
805623	10	29478204	Hs10_8882_34:15	11449523	none	intergenic		none	none	F	[A/G]	G
1019850	11	83985464	Hs11_34082_34:6	14458185	NM_001364	intron		DLG2	1740	R	[A/T]	A
1599931	12	56545117	Hs12_29578_34:1	20402156	none	intergenic		none	none	F	[A/G]	G
AA	13	29711866		NM_006644	exonic	intergenic		HSP105B	10808	F	[A/G]	G
279941	13	101396168	Hs13_10109_34:1	16787844	NM_000452	mRNA-utr		SIL10A2	6555	F	[T/G]	G

RS_ID	Chromosome	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
1062230	14	2347927	Hs14_26604_34:1	5409277	NM_014178	mma-utr		STXBP6	29091	F	[C/T]	C
185911	14	70531426	Hs14_26604_34:1	52461426	NM_004296	intron		RGS6	9628	R	[T/C]	T
1477261	14	76211019	Hs14_26604_34:1	56141019	NM_012245	intron		SNW1	22938	F	[T/A]	T
1191119	14	98323228	Hs14_26604_34:1	80262328	NM_004434	intron		EML1	2009	R	[G/A]	G
657780	15	60662497	Hs15_10351_34:1	33936998	none	intergenic		none	none	F	[A/G]	A
1393890	15	95036282	Hs15_10431_34:1	12272041	XM_096908	UTR		LOC145945	145945	R	[G/C]	C
1478714	16	7132995	Hs16_10709_34:1	4474136	NM_018723	intron		A2BP1	54715	R	[G/A]	G
888213	16	66996694	Hs16_10655_34:1	15942010	NM_178516	intron		LOC283849	283849	F	[C/T]	C
690115	17	73471192	Hs17_10798_34:1	6766230	NM_178128	mma-utr		LOC283985	283985	R	[G/A]	G
1465501	18	18066783	Hs18_11123_34:1	1303898	none	intergenic		none	none	F	[A/G]	G
899173	19	14094734	Hs19_11452_34:1	5496536	NM_018154	intron		ASF1B	55723	F	[C/T]	C
10477	20	38236897	Hs20_11519_34:8	2804383	NM_015568	mma-utr		PP1R16B	26051	F	[C/T]	T
926393	20	38331457	Hs20_11519_34:8	2698943	NM_021931	intron		DHX35	60625	F	[C/T]	C
465271	21	26507612	Hs21_11689_34:9	13247612	none	intergenic		none	none	F	[C/T]	C
13847	21	44259022	Hs21_11672_34:1	715558	NM_020132	locus-region		AGPAT3	56894	F	[A/G]	G
738658	22	30168847	Hs22_11677_34:9	11234862	NM_019843	intron		EIF4ENIF1	56478	F	[C/A]	C

[1] Contig Accession Number which can be found in the NCBI Database:  
http address: [www.ncbi.nlm.nih.gov/entrez/query.fcgi](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi)

[2] Sequence Identification or Nucleotide Accession Number which can be found in the NCBI Database:  
http address: [www.ncbi.nlm.nih.gov/entrez/query.fcgi](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi)

[3] "A" column is the sequence orientation ("F" is forward, "R" is reverse).

AA genomic sequence (SEQ ID NO: 4)

TCATTAGCTTTTTAGTTTTTCACATTCCTGATACAGACGTAGGAGTGCTCGTATTTT  
GGATTTTGCATCCAACCTGTACTTTAGTTTTAAATTTCTGCACA[A/G]AAATGTTCCACT  
AATCTTTACATCGAAGTTTTTTCTCTCTAAGAAAGGATCAAAAGCTGTCCCCAGTACC  
TAATTTGGTTGAAACAACAAATAGTCTGTT

[0255] Following are genomic nucleotide sequences for a *KIAA0296* region (SEQ ID NO: 1), a *chrom 4* region (SEQ ID NO: 2), and a *SNW1* region (SEQ ID NO: 3). The following nucleotide representations are used throughout: "A" or "a" is adenosine, adenine, or adenylic acid; "C" or "c" is cytidine, cytosine, or cytidylic acid; "G" or "g" is guanosine, guanine, or guanylic acid; "T" or "t" is thymidine, thymine, or thymidylic acid; and "I" or "i" is inosine, hypoxanthine, or inosinic acid. Exons are indicated in italicized lower case type, introns are depicted in normal text lower case type, and polymorphic sites are depicted in bold upper case type. SNPs are designated by the following convention: "R" represents A or G, "M" represents A or C; "W" represents A or T; "Y" represents C or T; "S" represents C or G; "K" represents G or T; "V" represents A, C or G; "H" represents A, C, or T; "D" represents A, G, or T; "B" represents C, G, or T; and "N" represents A, G, C, or T.

KIAA0296 genomic sequence (SEQ ID NO: 1)

>16:31076951-31174000

1	ccccaccccc	caacagctgc	acagctctgga	gcgaatatac	acgcccacca	cccacacacc
61	caagacccaa	tacacttttt	taaactttat	ttttacttct	attttatttat	tttttaattat
121	tttttaaaaa	tctaattaga	gatgaggtct	taggctgggc	acagtggctc	atgcctgtaa
181	ccccagcact	tcgggaggcc	gaggcaggca	gatcacgagg	cgggaggatc	acgaggtcag
241	gagttcRaga	ccagcctggc	caatatgtgt	aaacccatc	tctgtctaaa	atcacaaaaa
301	gagctggcgc	cggtggtgtg	caactgtaat	ctcagctact	tgggaggctg	aggcagaaat
361	gtttgaactc	aggaggcgga	tgctgcagtg	agctgagatc	gtgctcactg	ctccaggtct
421	gggagacaga	gcgagactac	gtctcaaaa	aaacaaacaa	acaaacaaca	acacaaaaaa
481	acagagataag	gtcttgccat	gttgcccagg	ctggtctcaa	gtcctgggct	caaaggatct
541	tcctgcctca	gcctcccaaa	gtgctaggat	tacaggcgct	aaccactgca	cccaccctac
601	tttttttttt	tttttttttt	atacaggatc	tcactctgtc	accgggctg	gagtgcaagt
661	gaagatcac	tgctgactgt	acccttgacc	tcagggactc	aagtgtacct	ctcgtcctcag
721	ctcctctgagt	agctgggaact	acaggagagc	gccagcacac	ctgggtaatt	aagatttttt
781	ttgtagagac	agacgtctat	ttgccaggcg	tgctctcgaa	ctcctggctt	caagtgatac
841	acccttgccc	tcctaaagt	ctgggatcac	aggcatgagc	cactgcacct	agcctaatat
901	agttaatatc	cccgtaagg	ctgctcagag	ggcctgagag	gaacaaaagg	ctcagctctg
961	gagagctcca	ccccagcgc	caatctctct	aaatggcctc	tttctctccc	atatcccaac
1021	caaggctctg	gagtcagatg	tcctgtgacc	ttaagtcaacc	attccaaagg	cctgcgactg
1081	ccccagagga	ccacaagtga	aataatatta	taactcctgag	aagttttagt	gcacagatgt
1141	gcctgccatt	aagacgtctga	gaaacaaaga	ggaagatggg	accaggggggc	ccagaagacg
1201	ctggaaacca	cagttattaaa	agctcagaga	ggctgggcac	agtggtctac	acctgttaac
1261	ccagcacttt	gggaggccaa	ggtgggtgga	tcacttgagc	ccagggggtt	gagaaacagg
1321	tgggcaacat	gcgcaaaccc	agctctctacc	aaaaaatata	caaaaattag	caggcagatg
1381	tggtgcgtgc	cttagtacca	gctacttggt	aggctgagcg	aggaggtatg	actgaacctg
1441	agagcacacc	actgcactcc	agcctggatg	acagaaacag	actgacctc	aaagagaaga
1501	aaaaaaataa	aaaaaaaagg	cccagagggg	agggTaccct	caacagtttt	ccagccccct
1561	ccacatctct	acataacctca	cttgatagtg	ttcaagtcct	accttgagga	ccagcagaat
1621	tataggacca	agccgccaaa	tggggaaatt	gagtcaccga	gagaagtaat	gcattattta
1681	agatcccatg	caggactatg	agtcaggggg	ccaagagacc	ttccaccgtg	tgccactcag

1741	agacacagag	taggaggggg	aaggggggtcg	ggtagggcagg	gacaaaagat	gcaggaggcca
1801	acgacagctg	actgaagagg	acgagggctga	catgaagaag	ccaggagcagc	agactctcttc
1861	cttatcatct	ccaggggaca	ccactgggca	gggcttgccc	tccggaaaaa	ccctgcattc
1921	ccctcttggt	ttcatcaggg	gacctctctc	ctactagtct	gggttttttt	gggtttttttg
1981	ttttgttttt	gagacagagt	cttactctgt	caactaggct	ggagtgcatt	ggcgtagact
2041	cagctctctg	taacctccac	cttccatgtt	caagcaatcc	tcctgtctca	gcttcccaag
2101	tagctgggag	tacaggcacc	tgccatcatg	cctggctaatt	ttttgtatgt	tgttagagac
2161	agggtttcgc	catgttgccc	aggctggctc	ccactcctg	gcttcctgga	attctgcctg
2221	ctcagccctc	caaaagtctg	ggatttacagg	catgagccac	cacaccctgc	ctgagctggg
2281	ttttaacagc	aagaggagaa	gagccaaaaa	tcctcacata	gaatcacaca	gcacttgaca
2341	gtttcccaacc	tcactcatcac	tgaagtttag	agcagccgat	accctaYaaq	atgatctccc
2401	catcccccta	cagttaccaca	ctgtgcagag	ggagatccac	acttagagac	agaaagcgt
2461	ttccagaaagt	ccatagcaac	tcagttccac	gaatctaggt	ttcctgacca	gggcatagac
2521	gaaagggtcc	attccttttc	ttgcttgatc	cttcacagaa	gcttcctgga	cagagccctg
2581	gggtccagga	gacctgttat	tcattccccg	ctatgctgag	acttgctgag	tgacctttgg
2641	gactctctct	agagaatata	agttccacga	gtgcaggaa	ttttgtctat	tagtcttgat
2701	tgtattctcca	gccttagaac	agtggttgcc	ccatactagt	tgccccaaaa	atatccattt
2761	aatgaactga	gtgtgctgtg	catgggtggt	catgcctgta	atccccagca	tttgggaagc
2821	tgaggcagaa	ggattgctta	agccccaggag	ttagagacca	gcttgaggcaa	catagtggag
2881	ccgcatctcg	taaaaatttt	taaaaataaa	aaatgagtga	atatcttaga	agccaggtat
2941	agagaaagtgt	ccagctcaga	aagcctgaag	cctaagaaga	accaaggaac	caggggcttt
3001	atccctcagat	acatgaaagc	ctgaattctc	gtccacaagt	atttatagag	ggcccgtaat
3061	gtctcttgta	ctgggctagg	aactccccac	attcagttaa	gaacaaaagt	attacctggc
3121	ctccagatga	aggcaggggc	tggggggtgt	gagtgccagg	gagggcagct	gatcaataca
3181	aaactttttc	ttagcctgag	ctgcccctgac	atggcttgac	ggcttcaaac	gtggtagtgt
3241	cagccgggct	cgagtgttca	aggagggcgc	cgccctggcc	cccaactctc	agccagctgc
3301	ccagaaagat	gcggaagaag	agatttctgc	cttggtgag	ctgacttccc	acccccagat
3361	tcctctccca	cacacccctg	caattttctg	acgtgcagca	ctgcgaatct	atctttcccc
3421	gattttcaag	gtcccatgat	gctgacagcc	ccaaaagtca	agtcgtcagt	ccggcccagc
3481	cctggaccgc	aaagcaataa	aggcgaggtc	agcaaaaggt	ctaccaccca	ctgcccagaa
3541	aggcctctgc	gggtggtcgg	cgcccccctc	cccaactcgc	ggggggccgt	tgggcgtcgc
3601	tcggtgtgtg	gggtgccggg	gacgtcgtga	tgagaacgca	gtcccagaga	gagccgtgtg
3661	agagccggga	cacgtgacag	tcacagggtc	acattctcgc	gtcccacagt	ttgggacggg
3721	gctgtgtcac	tgacgcggtg	ggggcaccat	gggtgagtgt	gagatcgagg	gtgtccctgt
3781	taactacaaa	tgacttatct	ctaccccttt	tgccaaccag	atttccgtgt	gaagatgcac
3841	cggttccggt	gacgttagca	agtttctcgc	tccaggcatc	tccgcttccg	ctcggggcgc
3901	aaacaacttc	gactccacct	tcccagcctc	gggcaaggaa	gagacgcgac	catgtgcgca
3961	tgccccgaat	ttatcacgga	ggggcggggc	tgaggctgcg	ggagctggag	cggggaagaa
4021	aaaggaaatc	caacctgtgg	aaccttgggg	ggtccccggg	gtcggcgcc	tccatttgac
4081	tgtagggcgt	gcaaggagac	gagcctctgg	cggtcgtggt	gggtgttggt	gtccgcagggt
4141	ggagggaggg	gagtgatcaga	gtgtgacggg	ggtagcgtaa	ttccaaaatt	ggaggcctcc
4201	cgctctctgc	gcgggggagc	gagagctcac	gcgcacatgc	ggagcaggag	cacagagctc
4261	agacaggtga	gacgccaagg	cagcggggat	ggggagcggc	ggaggaactg	gaacggcagca
4321	cttctgtctc	ctgggatagc	gaggggtggc	tgatggccag	aaggaagagt	cccggaaacc
4381	tgtaggtctc	gcggggttaag	agccgcagcg	aaacggtggt	gccaatgact	ccgggctcgg
4441	caagggaggt	aaacgctcga	cgaagaggag	aaaggaaggc	tcgcgtgact	gtgtcacccg
4501	ggcagcgac	ggctggggag	cccggacag	gagtttcttc	acaaggtaag	gggctggggt
4561	ctccgctggt	attcgcgagg	gtgtaggagg	accggaggag	aggtccacct	tgagatcccc
4621	catatctctc	tcagccctct	cggtcacctc	ccccaggctc	ggacaattcg	cgagcattat
4681	gtcaaaactg	ggaaataaagt	ccaggagttg	gagaaacagc	aggtcaacct	gctggccacg
4741	ccccctcccg	aggagagtga	gtgaaacccc	ggctgcaggg	cgcatgctcc	gtccccaggga
4801	ttgtgggggt	tgtagttcca	cgcaggtggt	ggccagagtg	gtttgttgag	gttggggctg
4861	ctgtttggga	gtcttggcct	tctcttatct	aggcatgaag	caggagctgc	agaaacctgc
4921	cgatgagatc	aaacagctgg	ggaggagagt	ccgcctgcag	ctgaaaggct	gtctcctggg
4981	acctcagaca	gtactctccc	tctgatctgt	ccctgttgtt	ggatatctgt	ggagagtgtg
5041	ggccccagaga	agccagtgat	atatccaggt	cacacagcag	gctcgggtct	agcatcgtc
5101	tcctggcctc	caggccattg	tactctccac	agcacaagtc	cgccctccag	gttcttttat
5161	ttacaatgaa	accatttact	tacacagtta	tcgtctccca	ctgggctcat	tttgggcagg
5221	gagatggagt	tttgttaggt	ggcctctgca	tacctatagt	aaactcagtg	tgtaatgcac
5281	agaaaaataa	acttactttc	tctctttaga	ggctcagcct	tagtcatgtt	atgataaatt
5341	atattttccct	aaaaatccta	tgagagacaa	tacccccact	acccttgtgt	gttccccacg
5401	ccatagagcc	ccagaaggag	gaagctgagt	agaaactata	ccgctctaac	acaaagatga
5461	gaaaaaccaca	gggtgggtttt	ttttctcaga	aatgaggaca	tttcagcaaa	tggtttcatga
5521	agtattagat	gacaggtgta	tgaagggaag	gcttcgacag	atcatggagat	ccaatttgat

5581 gacttttcca atgtgggaaa ctgagctcag agagagaaag aacttgetca aggtcaggaa  
5641 gccagggtctc ctgatgtctca gacaaagacct aacacctctgc tttattttct tccattccat  
5701 aggaagtttac tgtgacccca gacaaagacct agtcttggtc gtggggacaa tggttttctt  
5761 tctttttttc tctcagctctc ctgaatagct ggattacac gcggacaccc gcgtgcccag  
5821 ctaattttttg tagtttttagt agagagctggg ctccaccatg ttggccaggc tggtttcgaa  
5881 cctcctagctc caggtgaccc tctgctctcg gctcccaaa gtctgaatt ctgttgaa  
5941 gccaccatgcc ccagctggga cacatgtttt ctgggagtca agatggaggag ttaggggtca  
6001 ataggggata aagacattac tccagctggga cctgtgtggc acgcggctgc cccaggggag  
6061 gagagtgaga agtcataaat gactggcagg ttctctatct atgtgacag acatctctta  
6121 gtccccagca tgaatlcata gaagtcaaga agaggaacct ctgtggggca gacatgaaga  
6181 ggaactcccc tgggtgtgata tcttattttt ttaattatta tttatttttt gagatggagt  
6241 ctcaectctgt cctcagctgt ggagtagact ggcacaatct tggctccact gcaactccac  
6301 ctctctcaag cgattctctc gctccagctc cagcagtagc tgggattaca ggtgtgcacc  
6361 accacacctg gctaattttt tataattttt gtagaaatga gtgttcacct tgttgccact  
6421 gctggctcgc aactcctgaY ctcaactgat ccaactgctt tggcctcgca agagtctatg  
6481 attataggca tgaaccaccg cgcgcggccc ctgggtgtg atcttagtaa ggaatttgc  
6541 agtgatctga ctggccctct ctgggtcccc agtgaggagg ataccaggag gtcagggttg  
6601 gagtagttgc gccagggtct cagcagggac cccagattga gcttgaggca gcttggagga  
6661 ctgtgaaggag tgaagctgga accaggaagc agatgtatc tcttgaaaaa gacttccaa  
6721 gaatgagcat atttaaggcc tcagaaagag gggaaggcca gacgagatgc cccagaacca  
6781 gtgttttctg ggaagcctgt ggtggtgatg ggcattagtg gttagggctc catgtggccc  
6841 tgttgccact gtttcgcccc ggcaaatgt tcatcttag gcgttaggagc tgtggttgag  
6901 gcagcgaggt tggcattcag caagcattca gcagttaact attgggtgccc tcttctgtgc  
6961 cagacacttt tgaactgttt taggatacag cagtgaaacca gtgacctgt tactctagg  
7021 acttccctct tgggtgtagac aatcaccata ataaataagt gaattattga gaacataata  
7081 agcattaaag aaaaagagc aggggaagag ggactaaaga tgtttaggga tgcatttgga  
7141 cagttgaaggc caggtggagg aagcttctat cagaaggtaa catctgaaca agagacttaa  
7201 aggtgtttgc tgggaatgag catctaggt agaaggaaaa gtagctcgaa tagaggctga  
7261 tgagagtgtg ctttgtctag ggaggggtaa ggagaccagt ctggatgggc agaggaaggc  
7321 aacagtaaga ggaagtaaga tcagagaggt catgggagaa ggaagatca tagaggctga  
7381 gccagcgccc gtggctcagc cctgtaatcc cagcactttg gaggctgagg tgggagttt  
7441 ggttgtagcc aggagtttga gaccagcctg gccaataag tgagaccccc ccccctttat  
7501 tttttttctc ttgagacagg gtctcactct gttgcccgag ctggagtaca gtgttgatt  
7561 ctctgtctac tgcaacctcc gctcctggg ttcaagccat tctctgtctt agactccca  
7621 agtagctggg actcagcgcc cccaccactg caccagacta atttctgtac ttttagtaga  
7681 gatggggttt caccacgttg gccagctgtg tcttgacctc ctgacctcag gtgatccccc  
7741 tgccctcagc tcccaaaagt ctgggatac aggcattgagc caccctgccc tggaggccga  
7801 gtctctatta aaaaataaaaa tagggcagggt gcagttggct acgcctgtaa tggaggccga  
7861 ggcagggtga tcacaaggtc aagagatcaa gaccatctgt gccacatgta tgaaccccca  
7921 tctctactaa aatacaaaa attagccgtg cgtggtggcg cgtgctctga tctccagcta  
7981 ctccgggagc tgaggcaaga gaattgctgt aacccgggag gccagggttg ccaagtagcc  
8041 agatttggcc actgcactcc agcctgggca acaagagtga aactctgtct caaataaaca  
8101 ataataataa aataataaaa taaataataa aataaataaa taaaagaagt catggaggac  
8161 cacataagcc tgataaggcc tttggctttt agtctaaagc aaatggggga gctgtcgaag  
8221 gtcatcaca ggtggttaag gtggcagatc ccgcaataga gctcatgca tttgtcact  
8281 gtactatggg gttgccgagg caccgaccgg caggagatcc tcccaggggc actcagctca  
8341 tattcttcat ctttagcatg ggggtctctc ccagcaatc gtggagctca tcaacaagt  
8401 caatccaagt cagtcggaat accgggagaa gaaactggag gaactctgga ggcagctgaa  
8461 gatcagtgag ttgtgcatgc ccagcctggc ccgagggggc aggtaatccc aacccaaccc  
8521 tgagcctggc ttttctcttc acagccaagt ctgggattgt gtctgatgag gatttggagc  
8581 agatgtctgca cagtggggca agcagaggtg ttgtgtccaa tgtgagtggc cacagccagc  
8641 cctctctcgc tgtgcctcca atccctctgc agtctctgcc tttctctgac ctctctggct  
8701 cagctgatcc tcttgcctca gctcccgag taagctgggac tataggtgca agccactgca  
8761 cccgctgtgc tggggccctt tctgattaa ggcaacctga ggcctctaa ggcacttaatt  
8821 agcctgctgc gagtcaacca tcagattcca ggctgagggc tcccagaaga ctcaacaggc  
8881 gtttctgacc tgcgtcggt ctcctgtga acactctgca cactcctgag aacccccagc  
8941 atcttgaagg acacgcaggt gactcgagac gccttaagt agactctggc ccggccagat  
9001 gagatccagc agcttgaacg cagatattcgt gaggctccag acatatccac ttttctggct  
9061 accgaagtgc agatgcaggt ggggtgcccg cgcagcccca gacgtgagac caggtctcag  
9121 ccaaaactgcc agMctccgc caYccttga tttctctccc gaggtctttg gttcttcag  
9181 gtttgccatg gcccccagat tgggtcttat tctatctct agcgtgaccc cgagaaatgc  
9241 acctgctctc gctgctacac agatgccacc tccctctgct atagacacct gccccctcct  
9301 caaaaactgca gctgcccagc gttcgtgcccc agccctcact ccccctccac taacagatc  
9361 cacccttata cctctcagag tccagctcag agttgcccata gaggggctgc cctctaacat

9421 ctgtacaaag ctgggggtgg ggcggcgttc cctctggcct ggttgtgagt Wgaagtgagc  
9481 ttccagccgt gtccctggag agctggccct agtcagtca cagccaatgc ccttttgagc  
9541 ctgagaccta gcggaagag atctcattca gtaggagtac tagaacctga ggctgtgtgt  
9601 ccagagagga gccagggata gggagggtct tgacagagct tgagatagc ctgggaagat  
9661 gggtaaatct agacagattt gtgaagycac agttcccat ctgtgaagg taatgacat  
9721 ttgaggccct tagctccaaq ctaccactgc agatagaagt tgtatggagt aagtgagcag  
9781 ggacacaggg actacatgat agaaggggcc ttgaagccat ccccaaggag tctgaacctt  
9841 gtgtcagata agctcttgcc ttgtctttgt tagtgcatt ttttttctt ctgaagcaatg  
9901 ttctctagtc attctggggtg ggggtgggcaa aggcactcct accctccctga accaccocat  
9961 cctctgagca gggggagatg atcaatcgca ttgagaagca cactctcgca tgcagggact  
10021 acgtggagcg tgggacgagc cagctcaaga cggccctgga gaaccagaag aaggcgagga  
10081 aggtgagcct ccagggcccj gccactgcc caggcaccct gtgtgacttc cctgaccocc  
10141 tctctccca cagaagaaag tcttgattgc catctgtgtg tccatcaacc tctgtctctt  
10201 agcagctatc attggcgta cagtgggtgg ataagtgcg acattgttgg tgagatgttg  
10261 tgggctgcc cctggcctgc cccagccctg gccccagccc tccctctctc ctgacacctt  
10321 gttctccct ctttcttac aggcactagg agccacagga acccggggcc ttgctctctt  
10381 tccacagcag cttgggggca gggcagagcc tccagtcgga cccctctctc acactggccc  
10441 ctatcgagca gggcagacag ttcttctggg gttggcagct gctattccat gatggcctc  
10501 tctctcagg ctcaatgctt gggggaggcc tgcactgtcc tgattggccc gacacacagg  
10561 ttttgaataa aattaaaaa aaaaaaaaga gcatagaaag cctgtgcac gtgtgttact  
10621 ggaagggctg gcccaaggct tccgggctac caactcctt accctctgga cgtccccagg  
10681 gccaggtctg gccctggctg ctgagtcgaa actgccaggg gtgctgtgca cagcagagc  
10741 tggttctctg ttcttgacc cccatagaa ttgggtgggca gggaggggta acaccggcat  
10801 cttctctctg cctcagta ctccccggga aaggaccact gtgagtact gtcttggaaa  
10861 tgatgggctg gaccagggca ggctgggagc caggtgagat ggggggttgg gtggcatcag  
10921 tgggtctctt gttggcccag aggaagagcg accatgaaa aatgtctat gtgtccgaa  
10981 actttggagt cagtggatag ggatgtgtgt gtttcagcag ggatgacat ggaggtggat  
11041 gtttaagctg ggaagaggtt gccagtcaga aagcacagga acccggggccc ttgtaacca  
11101 aaaaactata tcttttacct aagcgtttag gcagggtgtg gtggctcgca ccagtaact  
11161 cagcacttcg ggagggccag cgaggagagt ctcttgagcc caggagcttg acagaccct  
11221 aggcaggagt gggcaacctc ttctctttag agaataataa ttttcaaaat taccaggggc  
11281 tggatgcaga ttgtgtgtgt cccagctact ccagaggtg aggtgggag atcgctgttg  
11341 ccagagagat taaggctgca gtgagccatg gtcactccac tgcaactccat cctgggtgac  
11401 agcagcgagc cctgtctcaa aaataatagc aatcatcacc agtagcagca gcagcgagcag  
11461 cagcagcata gagagccagt gatcctggat cagtgacact ggttgcctag ggtaacctg  
11521 ctgaagcagg tgggtggcagc agaaaagcct gacctctgat ttcttccata cctgactcg  
11581 aatccaagcc ctgactaaat ttcttttttt cttttttttt gagacagagt ctgttctctt  
11641 tgcccaggct ggagtgcagt ggcactatct cagctcactg caagctccgt ctcccaggtt  
11701 cagcccatct tcttgccctca gccctccagc tagctgggag tgcaggcacc gcgccaccac  
11761 cccggtctat tttttgtatt tttagtagag acggggttct accgtgttag gatggctgtg  
11821 atctactgac atcgtgatct gccctcctgc gccctcccaa ggttgtggag tacaggcctg  
11881 agccaccgct taaatttcta agggctccta gtcctgatgc ctaatttctg tggtagagt  
11941 ggctctgtct ccccgacacc tagagttttt gtttgtttgt ttgtttttgt tgagacagag  
12001 tctcctctgt tgcgccagcc tgggggtgca tggcgcaatc tggctcact cgaagctccg  
12061 cctcccgagt tcaagccatt ctctcgctcc agcctccaga gttagctggga ctacaggcg  
12121 ccgcccacat gcccggttaa tttttttttt tttttttttt gtagcttttg ctctatgcc  
12181 cagagctgag tgcagtggtg cgatctccgc tcaactgaaa cctcgcctcc cgggttccag  
12241 ccatctcct cctcaggct cctgagtagc tgggaactaca ggcaccgccg acccgccccc  
12301 gctaattttt tgtattttta gttagagacg ggtttcatcg tgttagccag gatggctcg  
12361 attctcctgac ctctgcatcc gcccgctca gcttcccaaa gtctgggat taagactcg  
12421 agccaccggc cccggccccc gacacctagt tttaaaggtt aagccggctc ctggccacctg  
12481 cctacttgca gtaggggcggc gcttagctct gacctccaag gctctgggag tgggtcagct  
12541 ccgcccagtc catcccactt tcaatcttca aggcctctgc tgttgcctgc gctcccgctc  
12601 ctcccagctg cccagctctg cgggctcaat ccccgcttgc catgtgtgag catgtgtgag  
12661 cgctaagcg tggatgtggc ttctgtgatg cacatggag cgggctctgg accgtgctag  
12721 ggaagggcca gaggggcgga attggggccc agggccaggg ctgcgcagc cctgagctg  
12781 cctcccggtg gccccgcagc gccctccggg ggccctggag tgcaggtctt accgtccag  
12841 atcgttcgca actgggcgag ctgtgcatgg ggcctggcta agggctgct agggctacga  
12901 ttggccagcg ggaacttaag gttgtctctg aagagcatgg acattagtct ggaggtgtct  
12961 ggaagagtga tccccgcccc accatcaaat ggcctctagg tctaggaaag tgggtgtgtg  
13021 ggggccttag ggcgagggcg agacataccc cgaagtgtgt ggattgtata ccgcaaggg  
13081 ctggatcgaa ccccccaaa acactggag cgtgtgtggc ttagagaggg ccgacagat  
13141 cagtgtgtcS tgggctttac aggaagagag tccacctct ctgagagtgc cagatgcgat  
13201 ctagggtgtg ccacccgatg ggagctgcgg gccgggcaga tctgccccca gtacaaagct

13261 gatttggaac tggggcctct ggacttccct gattctctgc ttgcatctcc agcaaatgcc  
13321 tgtccctgtg tgcctcttca tccactctct cacttctctg tcttcagagt aaattgtcaa  
13381 gatctgttgt gcttactggg atctgtataga gtctctcggc atctcactgtc tatgcagcgg  
13441 tggctccact gcagcggggg ccagtgtgac cggggggcca cgtgcaggtg tgcctctctc  
13501 ttgagccatg tggacagcgc cgccctgtga aagcagctcc ccggtttcac ccagaaagcc  
13561 atccagaacc tggcagaaaa ggtggcctga tggccaagtg ctgtcggagt caggctctca  
13621 tccctttgaac ttttctgtgt ggctgtcagg accatagaaa ggtctttgag caggtgtagt  
13681 tggagcagat tgggtaggca agcgaacaga tgggtYgt cactgtgagt tccgttgggt  
13741 cccctctgtta catctcttcc ctltgggaaa ctgcccgtgag tggagggcta agggcaggat  
13801 ttgcatgtga atctttagct tgtctgtctc agcccaactt ttagccaaca ggttctrtgt  
13861 ttgagtgtac atttccaaat ccatcttgtta tcaaaactgt tcaagtgcag ctactatt  
13921 caatcattgt tggttcaagt tcccaagaaa atgaatcagt ctggtctcgt cccagatac  
13981 gattacgttt acttgcctag gaattgtctg cccctttaat caaagctttg cactgttgtt  
14041 cacatttgtta atccacagcac ttgtggagcg caaagcagga gtattgtctg agcccgagg  
14101 ttcaagacca gccagggaaa tataacaaga cccatctctc acaaaaaata aaatttaggt  
14161 gggcctgtg gctcatgctt gtaatcccag cactttggga gttcctgcca ggtgtgtcac  
14221 ctgagtctgag cgttctcgaa ccagcctgac caacatggtg aaacccctgc tctactaat  
14281 acaaaaagtt agctggatat ggtggtgacg gctctgtact ctacttggga ggtgtgagg  
14341 gaagaatcac ttgaacccgg gaggtggagg ttgcagttag ccagagattg cccattgac  
14401 tcccaactgt gcaacaagag caaaactccc tctcaaaaaa aaaaaaaata aaaaaagcc  
14461 caggtRctg tcaagtgttac gtgctgtgtg tcccagctac ttggggagct gaggtgtgag  
14521 gattgtctgt cactggggtt gagaccacag tgaaccaata tgcaccact cccgtcggc  
14581 ctggacacaa gaataatacc ctgtctcaaa aaaaaaaata aaaaaaaata aaaaaagaa  
14641 aaaaaaaga ctttgccttt gagtcaagac ttacccttgg taaccttggc gaagtgtgag  
14701 gttaggaagt acatgttaca aaatgtgtga gcaagcgtgt tgtattgtct ggaagaggag  
14761 ttgactaggg cagtgattga catctctgtt ccagataatt gcttaccttc agctgtgag  
14821 cccctccatc aggagcattta tatgtctatt cccacttacc aataggtttg gctataggac  
14881 ttgctttggc cagtggaata tgggtaggaa ggcacaaatg ccgctggggc gatgtgctca  
14941 cactcttaat cccagcactt tggggaggtg agggcgggtg atcagctgag gtcaagattg  
15001 cgagaccagt ctggccaacg ttgtgaaacc tgtctctcac taaaagtaca aaattagcca  
15061 ggcattgtgtg cagcggcctg taatcccagc tgcagggaag gctgagggac gagaatcact  
15121 tgaacccggg aaggaggagg ttgcaatgag cccagatcat gcatgtgac tgcattgcca  
15181 acaaaaagtt aaactccgtc tcaaaaaaaa aaaaaaaagt taaagtctca ctctgcata  
15241 gaagctttag ggcacacttg agtactctag cagcttccag tctctctccc tctgtcag  
15301 gctcataaac ctggcagttt ccagatctag acttctcttt cagcctgcga cccagaatga  
15361 caatgacatg aactgtggct acagcctacc tataaaatga tgcagaattt aagaataat  
15421 tctctctctg tgtgagccat tgatatatgg aggttgtttt ttagcacatc caaatgttta  
15481 aacaaactgt tacagaat ta ataccagaa gtgtgtgtgt atcaaaaata aaattgagcc  
15541 tcagccggcg acggtggctc acacctgtaa tcccagcatt ttgggaggcc aaggtaaagt  
15601 ggtcacctaa ggttaggagt ttgaaccag cctggccaac atgacaaaac cctgtctcta  
15661 ctaaaaatac aaaaaaatt agccaggcat ggtgttaggt gctgttaact ccagctgtcta  
15721 cgtggcgagc gagaatctgt tggaccagg aggcagaggt tggcagtMg tcaagattgc  
15781 gccactgcac tccagcctgt gcgatggagt gagactccat ctcaaaaaat taataataa  
15841 aataaaaaat attattaaaa attagccagg tgtgtggca tctgctcta gctccagct  
15901 cttgggagcg tgagatggga aggtcacttg agcccaggaa gcagaggttg cagtaggcca  
15961 agattgcacc actgcactcc agcctgggtc agcatctctc aaaaaaaata cccagctcatg  
16021 tggctctact cctgttaatc cagcactttg ggaggctgag gtgggtgtgt aggtcaggag  
16081 ttgtactact gctgggcaaa catggtgaaa cctgtctccc actaaaaata cctccagcct  
16141 ccaggcatag tgggtggcac ctgtaatccc agctactcag gaggtctgag ctggaatc  
16201 ccttgaacct tggatgcgga ggttgcaagt agccaagatc aagccatgc cctccagcct  
16261 gggcgacaga gcaagactat ctcaaaaaaa aaaaaaaagt cctaactact gtactgaa  
16321 tgcattgac gttgagcttg acagtggctg gtaagggaac tgtcattgga atctgtaa  
16381 atggtgactg gtgttatgca atggtgaatc gtttgggttaa actgtgaagt tatgaccaa  
16441 tgaactttag ccttttagta aagaactggg aaaaaggagat atgtgtagca tgcgtctac  
16501 actattgcat gcaattgagg agttactaga agaaagagat gactcagaaa ttaaatgttc  
16561 agtttataag cagaaatgga agagaatata gaaattcgag gcaagtgatc tcaattgtc  
16621 gtaaaagata caactgagaa agtccctgag ccacaggtt ttcgtttttt tttttgagc  
16681 agtcttgcct ttgtttccaa ggccaccttc tgggttcaag cctttctctc cactcagcct  
16741 cccaagttag tgggataaca ggcgtgcacc gctaattttt gattttttag caaatgtc  
16801 tagagacagg ttccacctg ttggccaagc tggctttgaa tgcgtacct caaatgagc  
16861 tcccaactcR gctcccaaaa gtgcgtggat tgcaggtgtg agcagctgag accggctgag  
16921 tcaacagttt tgatataaag tcatctttgt ggcaagggcc atataagta tatgttgtt  
16981 atgccccttg taataacttc caaactgatc aaagtgttct ctaataaact ctctcagta  
17041 tcaagatga tcaaaaggaa agaggttaag agtgttaact cctctgctg ggcgtgtgtg

17101 ctcaagcctg taattccagc actttgggag gctgagggtg gcggatcacc tgaagtcagg  
17161 agttttgagc tagcctaacc aacatggaga aaccccgctc ctactaaaaa tacaaaaata  
17221 ccaggagcatg tgggtgcatg cctgtaatcc caactactgt gagggtcgag ccagaggaat  
17281 tgccttgaaac tggggaggcgg aggttgcagt gagccaagat caccactggc actcaggcct  
17341 gggccaagac agtgaacttc catctcaaaa aaaaaaaaat atglagctta cctgagggtg  
17401 ctacagtagct caactacagt taagtctaac gtcaatggta tgcttgaaaa gaattatggg  
17461 taagctgtgt acccatggat ctgaatggag taaaatacgt aagttcagtt ttggaggaaa  
17521 ttgcccctgct tcccctgctc aacacccctc caccctgaca aaaaagccac aggttgaatc  
17581 ttgaccaatga gtgttcaata ctactctgtc acccagcgtg gagtgcagtg ctgaacacac  
17641 tttattttcgg agaccgggtc ttctgggctc accagatcct ttgtgctcag atctccaaag  
17701 agctctctat aacctcgaaac gccaccggc caggctaat ttgttttttc aaattttttt  
17761 agctgggact acaggcccat caaccggcgt gaagtgcagt ggcacgatct tggctcactg  
17821 caaacaaaaa ctcaactctgc tgagtgatcc acttacctca gctccccaag tagctgggac  
17881 taacagttgt cctacacatg cccggctaat tttgtattt tttgtgtaga tagctgggac  
17941 tgcataattg cccaggctgg tctcgaaact ctgaactcaa gcgatccac tgcctgggac  
18001 tgcataattg cccaggctgg tctcgaaact ctgaactcaa gcgatccac tgcctgggac  
18061 tcccaaatgt ctggcaattt agggctgca tgtaaccaca tggccagcct attttttgtt  
18121 tgtttgtctt gtgtttgttt tgaagtgaag tctgtctgt tctccagc tggagtgcaag  
18181 tggcaacaac aagctcact cgaactctca cctctagggc tcaactgtac ccccacactc  
18241 agcctttctga gtagctggga ctacaggcat gcgccacac acctgtgttaa tttttctatt  
18301 tttttctgga gaggatttca gctctgtgcc caagctggtc tgaactctt ggtctaaag  
18361 agtctccctc gtttggcttc ccaagtgat gggattacag ttgttagcca catctggcca  
18421 cctaattttt gtatttcaag tttttttttt tttttttttt tttttttttt agtagagatg  
18481 ggggtctctg tatgttgcgc aggtcggcct caagcaatcc tgcctcaag tagctctctc  
18541 gccYagcct ctcaaaaatc tgtgattgca gatgtgaacc accatgcccg gcttgggtct  
18601 ccaaaattct tttttttttt tagagacgga gctcgttct tgaaccagg cctgagtgca  
18661 gtgtgtgtag ctgcgtctac tgaacgctct gccctccagg ttaaccggct tctcccgctc  
18721 caagctcccg agtagctggg actacaggYg ccgccacaca tggccggcta attttttttt  
18781 gtatttttag tagagacagg gtttctactg gttcggcagg atggttttga tctctgacc  
18841 tctgtatctg cccgctctcg cctcccaaa tggtgggttt atagggtga gccaccgacc  
18901 ctggccatgg gtctccaaat ttctatgggc atgaaggaga ctgagaaagc tactctactt  
18961 cagaaagaca taaccacag gtgctctca atgtggcca aggaataaa ttggaagaaag  
19021 tgggttttact ctaaggccag agccaagaac atgttgaaga atgaactagg gaaactcttc  
19081 cactccagg gaaaagtggg ggttctctc aacatctgcc caKcagcat ttagactatt  
19141 tggggccagg agcctgctgt gtgtctcctg tcttctcttc cttttttttt tttttttttt  
19201 tttttgagac agagtttcac tttttgtcac catgttttag tgaatggca ctatctggc  
19261 tcactgcaac ctctgctcc tgggttcaag cgaattctct gcctcagcct cccgagtagc  
19321 tgggaactaca ggtgcatgcc accacgctcg gctaattttt tttttttttt tttttttttt  
19381 ttgtttttga gacggagtct tgcactgtcg cccaggctgg atgtgaaagg ccgatctctg  
19441 gctcaactga acctctgct cctgggttca agcattctc ctctctcagc ctcttagta  
19501 gctgggacta caggggcccg ccaccaagcc cggttaactt ttgtattttt tagtagagac  
19561 cgggttttca tatgttggcc tggctggctc tgaactcctg acctgtgat ctgcctccct  
19621 cggcctccca aagtgtcggg attacagacg tagaccactg cgctcgctca atttttgtat  
19681 ttttagtaga gacagggttt caccgtgttg gccagggttg tctcaaacct ctgacctctg  
19741 gtgactccac ggctcggcc tcccaaatg ctgggattac aggcactgac ccgcgcagc  
19801 ggccctgtcc tctctctga acgggagtg ctgtctgtgt tctctgtctc ttgttctgt  
19861 ttatatgtgt gatgtgttgc tgtgtgtgtg tgtgtagaaa tggggccagc gtaactgtga  
19921 ctgtctctc tcttatttt tagctcatag gtctctgaat caagagaagc ctacatctga  
19981 cctgatatag aagagactat tagagatctt gggcttgagg ctgattccat ctcatagtg  
20041 tcaacttaggt ggtctccctt gggaaaggga tgactttatt ttgcatatgg aagaaaaatg  
20101 aaaggcagta ttgttaagga agaggggcaga cgggggaaag ttttataatt gttcaaaaag  
20161 attcaactgg atgtgtgtgg tggctcacgc ctataatccc agtgcttttg gagggtgaag  
20221 caggagatct acttgaggcc aggaatttga gaccagcatg ggcaacaatg tagagacctt  
20281 tctctcaaaa aataaaaaa ataaaaaaa attagctggg cgtgtgttggt ctgtctgtga  
20341 tctcagtagta cttaggaggg tagagtgagg ggaactacta agctcaggag tagagagctg  
20401 cagtgagtaga tttagtcacc atgcacatc gcactccagc ctgggcaaca gacaaaaa  
20461 ctgactctaa aaaaaaaccc aacaaaaaaa aatcacatgt atactctggc tctctctttg  
20521 gggacctgct acatagaatg gtttttttgt cccagttcac ttgacatcag tagagcttag  
20581 tggctgtctt tagaccatgg acttttagtg gaatagcat gtgccaactc cagcaggaaag  
20641 ctttaaaagc gctcatgggg tctgccaact ttccctctct cgtgtgtctg acagcgaaag  
20701 tccagccttg agacttttcc ttacagacgg gctctRgaat gaagatagca tagaaacag  
20761 tggctccatg gaggacatg atatgagtga gaaatcaaca tgggtgttga agccctcaag  
20821 atttgggggc tgcattact cagcagtaac tggatccagg ctgatagatg agcctccct  
20881 gtgggatacc ctgctcagg tctctttccc atccactga caactgcac accataatga



20941	gctatgctga	tgtaggaag	tctccgctt	tgtctctctt	cagaqctgtt	caccgtcagg
21001	tcttaacca	tgagctctt	tctttttttt	tctttttYtt	tgtttYttYt	cgagaWggag
21061	tcttctctct	tcaccagctt	ggagtgcaat	gggtcgatct	cgagtaaatg	caacctctgc
21121	cttctctgatt	tgagcaaat	tgtgtcttca	gctccttgag	taggtctgga	actcctgaac
21181	tcaggccatc	cgccagcttt	ggccttctaa	agtgctggga	ttacaggcat	gaaccaccgt
21241	gcaggcccaa	gcgagtgctt	cttgattctt	gctggcattt	cgactcagt	agacgtctgt
21301	cacaggaact	gtaaaaaat	ctgtgtgggc	ccagaccctc	tagcatcaac	atggtgccta
21361	gtaaaatac	atctcacatg	catcctgaga	tgcatataaa	agaaagcttc	agagccgggc
21421	acgggggctc	accgctgtaa	tccagcagct	ttggggagga	gaggcggtgt	gattgtctga
21481	gcggcgaggt	ttgagaccag	tctaggaaac	atggcaaaat	gtgctctcta	tttttaaaaa
21541	gggggggaaa	aagaaaaaaa	aaagctgggc	atggtggttc	acacctgtaa	tcccaaacct
21601	ttgggtggct	gaggcaggtg	gatcaacttg	gagaccagct	tggtcaaac	catgaaaccc
21661	catctctact	aaaaatacaa	aaattagcta	cacctcatgg	tgcaacNcctg	tagtcccccc
21721	tactcgagag	gtgagggcag	gagaatcgct	tgaaacctggg	aggtggagggt	ccagtgagc
21781	ccagatcagc	ccactgcact	ctaacctggg	ctagagagtg	agactctgaa	aaaaaaaaaa
21841	aaaaaaaaaa	gagaaaaaaa	cataatgttt	ggccaggcat	ggtgccttacc	actgttaact
21901	ccagcaggtt	gggaaagcga	ggggggcgat	caactgaggt	tagtccaaga	ccaacctaat
21961	caactgagtg	aaacctatat	ctactaaaaa	aaaaaaaaaa	attagccagg	cgtggtggtg
22021	gatgcctgaa	atccagccta	cttggggagg	tgaggcagaa	gaattgtttg	aaacctggag
22081	gcagaggttg	cagtgaaaccg	agattgtgtc	actgcactcc	ctgtccggcg	acaagagtaa
22141	aactccgtct	caaaaaaaaa	acaaaaaaaa	aaagaatact	aattgttagt	taagtgaatat
22201	tctgtaattag	tttgtgtatg	tgtattgttg	catataatag	agaccctaat	aaagctgggc
22261	tttaaataaga	tagaagttaa	tttctctctt	ctataaaagt	ccaaagttagt	atgattgattc
22321	tttccatgaa	atcattagga	gccagatttt	ttgtatcatt	ctattcatta	ttgttctact
22381	actaccattc	atagagacaa	ttttctgcac	catctaggct	ggagtgcaat	gggtgcaatc
22441	taattcactg	taacctcaaa	atcctgggct	ccagcgatct	tccctgacta	gccccaacaa
22501	agtagcagag	actacaagca	catgccacca	cgcttggtcta	atttttcttt	tctttttttg
22561	tagagtggtg	gtgttactat	gttgccagg	ctggtctcca	actcctggcc	tcaagttagt
22621	ctctctgctt	accctcccaa	agctctggga	tgacaggcat	gagcctaact	cgccctccag
22681	gtcttttttt	tcttgtgtgt	gttccatccc	tagggcggtg	ccctcaccca	catgatccaa
22741	tatgattcac	caccacttcc	acagctctgg	ccctctgagg	gggtgagtgt	tgccctttgc
22801	ctcaagagc	atgattcaga	agtacagatc	atttttgtct	tataccccat	agccagagtg
22861	tagtcatatg	gtcatatccc	gatgaaagt	ttgctgagaa	atagaaatctc	taccctgagc
22921	agcttttttg	ccagataaaa	gttcagttac	tctggggaaa	gggtagaagt	tgactgggg
22981	gaccataaag	tggtgccacc	acacacattg	aatgttaacc	catcccaact	gtatcaattt
23041	ttccttctct	tcttctcttc	ctccctctct	ccctccctct	ctccctctct	ctctctcttc
23101	ttccttctct	cttctctctc	tttgttctct	tcttttgaca	gtctccctct	atccccagg
23161	ctggagtgcr	gtgttgccat	ctcggtctac	tgcaacctct	gctccctagg	ttcaagcaat
23221	tctcctgcct	cagcctcctg	agttagctgg	attacaggcg	tgctccacaa	tgccccagct
23281	atttttgtat	tttttagtag	gacaggattt	cccatgttg	ggcaggctgg	tcttgaaact
23341	ctgccccatg	gtgatccacc	cacctcagcc	tccaaaagt	ctgggattat	atgctctgagc
23401	cactgccttg	gctcacaacg	gtatcaattt	tctgttactg	atttaaccaa	tattcatcaa
23461	ctcagtggtt	taaaaccaca	caactttact	ttcttcagat	tctggaagtc	agaaagttca
23521	aatcagttct	attgagccaa	gtctctgtgt	cagcagggtc	gggttttttt	gtgggtcttg
23581	gtggacaatt	tttcttgccc	ttcttcagct	cttttttttt	tttttttttt	tgagacacga
23641	ctcgtctctg	ttaccaggc	tggggtgcag	gtgtgcactc	atagctcact	gcagcctcca
23701	tctccccagg	tcaggcgatc	ctcccggtt	agccttctca	gtagctggga	ccacaggctc
23761	acgcccaccg	gcctctgtaa	ttttgtttat	ttttgtaga	gtagaggtct	cactccattg
23821	cccagactgt	tctcaaatgc	ctggattcag	gagatcctcc	tgctccagcc	tcccaaatgt
23881	ctgggattac	aggtgtgagc	cgltgcaccc	cacctctttt	cagttttagaa	aggtcactct
23941	tattccttgc	ctggtgggtc	catctccat	tggaagcac	atgaattcat	ctctgccttc
24001	atcatcactc	caactttctc	tctgagactt	atctcctctg	ttgtgcctct	gagtagagtg
24061	tcattgattac	ataccgcccc	cttggaataat	ctcgaataat	ctctccatct	caggatccct
24121	caacttttca	aaatcccttt	caccatataa	ogtgacattc	acagatttca	ggaaataggac
24181	gtagacatat	ttaggggggt	tctctattca	gctactgta	ccatgcattt	ccacacttaa
24241	ctctctcact	catttattca	taaaatattg	attgagcaag	acctgtgtgc	gagcacttgt
24301	gttaggttgt	agagaaatag	aggtgaaaa	acagacaagg	cctctgcttt	catggagttt
24361	atttctcgt	gaagaggaca	agtaaatagc	taagctatct	tttttttttt	tttttttttt
24421	tttgagagag	agtcctccct	tgctgcccag	gctggagtg	agtgccgcaa	tctcggttca
24481	ctgcaaggcc	cacctctgg	gttcacggca	ttctcctgce	tcaagctctt	gagtagctgg
24541	gactacagag	gcccgcaccc	acgcccagct	aaatttttgt	attttttagta	gagcaggggt
24601	ttccacgtgt	tagccagagtg	ggctctgac	tctcgacctc	atgatccacc	gtgctggccc
24661	tcccaaatgt	ctgggattat	agcgtgagc	caccatcccc	ggccaagagc	taagctattc
24721	taagctataa	cgtgtattat	caaaaacatt	aaggccaggc	acgtgtctc	acacctgtaa

24781	tcacaacact	ttgggaggct	gaggcggggt	gatcatttga	ggtcaggaggt	tggagaccag
24841	ctcgggcaac	atggttaaac	ccctgtctct	ctaaaaatac	aaaaaattgt	ttcgaggtgtg
24901	gtgtgtcagt	ccctgcagtc	cggctactcg	ggaggtcgag	gcacaagaat	agaattgtgt
24961	tgagtgggga	gggtggaggt	gcagtgaagc	aaagtcatgc	cactgcacta	caggctaggga
25021	gcacagWga	gcacctgtct	taaaaaaaaa	gcaatttagc	caagtgcagt	ggctctatgc
25081	tgtaatccca	ggcctttggg	agcccaaggga	ggggagatca	cgaggtcaag	aaatcgagac
25141	cagctgtccc	aaactgtgtg	acccctgtct	ctactaaaaa	tacgaaaaat	agctgggtgt
25201	gggtggcgct	ggctgtagt	ccagctactc	gggaggtctga	ggcaggagaa	ttgtgtgaac
25261	ccgggaagtg	gggttgag	tgagcccgga	tcacgccact	gcactccagc	ctcgacgacg
25321	agtgggaatc	catctaaaaa	aaagaagaaa	aaaatttggt	ggagatctgc	ttggaccag
25381	gggtggaggt	tgccatgagc	tgagatttgt	ccaactgcat	ccagcctagg	caacgaagcg
25441	aaaactccgt	ctcaaaaaaa	aaaaaaaaaa	tcccagcact	ttggggagcg	aggagggcca
25501	gatcacgagg	tcaagaatac	gagaccagcc	tgcccacact	gggtgaaacc	tgctctact
25561	aaaaatacaa	aaaattagct	gggtgtgtgt	gcgggtgcct	gatgtcccag	ctactttggga
25621	ggctgaggcca	ggagaaatgc	atgaacctgg	gagggcgagc	ttgcagtgag	ccgagatcac
25681	acagctgcac	tcacgcctgg	gcaacagagc	aaagtctgt	ttggggagcg	actggaggga
25741	gaaagcaaaa	gaaattaaac	agtgtagtgt	gacaaaaagt	gatagggggt	tggagacagc
25801	ttttctgtgt	gatgggttag	aatggctctt	tagaaaaagt	gactgcacaca	tgggagctta
25861	atgttgcagga	tcacgaggtg	aggagatcaa	gacctctctg	gctaacaacg	tgaaccccgc
25921	tccttactaca	aaaaatagaaa	aaattagccg	gggtgtgtgt	cgggcgctcg	gactccagcg
25981	tactaaaggag	gctgagggcag	gagaaatggc	tgaacccggg	aggcgagcgt	tgcatgtgagc
26041	tgagatcacg	ccactgcact	ccagcctgga	cgacagagcg	agactccatc	ctcaaaaaaa
26101	aaaaagataa	aaaagatggc	tgacacagag	ggcagagctg	agagcccaaga	gggcaaaaaa
26161	gagccagaga	aaaccatttc	caggcctgga	agcctaaagg	aatltcccag	ctcgatttgc
26221	agtttgcctg	gatttggtag	tcctttttac	ctttcttgt	taggggagct	gcaggttctc
26281	ttgtgctgtg	tgacgtacaa	gctccattac	accaagacaa	tagggatcca	cgacagagag
26341	ttacttgtctg	agggacacta	gtgcagagat	gggaagagcg	cctcaaatct	atctcccoga
26401	gcaattctcg	gagaggggtt	ctaaggggag	tgtggagggg	aggggaaatg	ggaggggtta
26461	gtttttgggca	actgggtcat	tgattgattg	ggggaaggat	gtgaagctcg	cgtttttggg
26521	ggaatttagct	ccctgtgggg	tccttcagggt	cgactgagtc	agtagttcca	tggagacctg
26581	aaggaatctc	ttttcttttc	ttctttcttc	ttcttttttc	tttttttttc	gagatggagt
26641	ctctctctctg	gcgcaggcta	gaggtgcagg	gggtgcagg	ctagaggtgc	ctaggtcaga
26701	ctcttggtcca	ctgcaacctc	cacctcccg	gttcaagcaa	ttctctcgct	tcagcctccc
26761	aagttagctcg	gactagaggt	gcgtgccacc	acaccagct	aatlttttga	tttttagtag
26821	agacaggggtt	tcaccatgtt	ggccagggtg	gtctcgatct	cttgactctg	tgatcgccgc
26881	ccgcgccacc	ctcgccctcc	caaatgtctg	ggatcacagg	agtgagccag	gtgcccagc
26941	cttaattttt	gtattttcag	tggagacggg	gtttcaccat	gttgatcagg	ctggagtcca
27001	atggttgcaat	cttggtctac	tgcaacattc	gcctcctgga	tttgaatgat	ctctcctgct
27061	cagcctccca	agtaactggg	attacaggaa	tgcttcacca	cgcccgcgta	atttttgtatt
27121	tttttagtag	agacggggtt	tcaccatgtt	ggtcaggctg	tcttgaacct	ctgaccttga
27181	gtgatccacc	tgctttggcc	tcocagagtc	tgaaggaata	tctcaaaaagg	aacacttaat
27241	gtttgtgtaat	gtccagggtt	tgatccatag	agcagtataa	ggtaaaagtg	actataaatt
27301	tttttttttt	tttttagaca	gagtcctctc	ctctgcacc	caggctggag	tgcatgttca
27361	cgatctcgcc	gcactgcaac	ctccgctccc	ctggttcaac	caattctctc	ctctcagcct
27421	ctcaagtgtg	tgctgccatg	ccagggtaat	tttttttttt	agacggagtc	tgctctgtct
27481	accagagctg	gagtgcaagt	gcaacaatct	tcgctcactgc	aaacctcgcc	ctctgggctc
27541	aaacaatgtt	ttttttcccc	tagtactttg	gtgtttgatt	atcttttttt	tttttttttt
27601	tttttttttt	agatagagtc	tcgctctgtc	accagagctg	ttgtgcagtg	tgccaattgt
27661	agctcactgc	aagctctgcc	tcctgggttc	atcccatctc	cctgcctcag	ctctccagtg
27721	agctgggaact	acaggccacc	accaccagc	ccggctaatt	tttttgtatt	tttttagaga
27781	cggggtttca	ccatgttagc	caggatggct	tcgatctctc	gacctcatga	ctgtcccgcc
27841	tcagcctccc	aaagtgtctg	gattacagcg	Rtaggcactc	gcacccagcg	tctgttgagc
27901	ttatctatta	tgtaaaacag	gctgggttta	gtggctcagc	cctgtaatcc	cagcaacttg
27961	ggagactgag	gtgggaggt	caactgagcc	caggagctga	agaccacgct	agcaattgag
28021	caactccctc	cctctacaaa	aagttaaaaa	atttagctgg	gtgcaccagY	agaccagctc
28081	ctcaggaagc	tgaggaggga	ggaactactc	agccaggag	tcaagagctg	cagtgagctc
28141	tgatcatgcc	actgtactcc	agccaggcca	atggagcgag	accctgtctc	aaaaataaat
28201	aaacatgaag	aagtgcgaac	acattatctg	gtttttgttt	ttgttttttt	tttttagagt
28261	gtgtctctgc	ctctcaaccc	tggtggaggt	gtagtgtgtg	gatctcgctg	actgcaaccc
28321	ctcgtctccc	gggttcaaac	gattctcccg	cctcagcctc	ccgagttagct	gggactacag
28381	gtacgtgcca	ccatgcctgg	ctaatttttg	tttttttagt	tttttttagt	gggacagggg
28441	ttcgtccagt	tgccagggtc	gttctgaaac	agatgaccca	ccacactctg	ccacactctg
28501	cctcccaaat	tgctgggatt	acaggtgtga	gccaatgtgc	ccggctgtgt	ctaaaaaacg
28561	atattgggcc	aactcggttg	ctcatgcctg	taattccagc	actttgggaa	gccaaagcag

28621 gaggattggt tgagcttagg agtttgagac cattctgggc aacctgggtg aacctgtgtc  
28681 ctgcacaaa atagaaaaat ttgccacctg tgctgtgtg tgctctgagt ccagcttact  
28741 ctcaagctgt agggaggagg atgtctgtg agctctggga agtccgagct cgagtgagcc  
28801 atgatcacac caccacatct tagctgcaca gaatgaagcc ttatcccaaa agaaaaata  
28861 aatgatattg tattatattg gaactttgaa ttatatgtg taactatctga agtttgaaat  
28921 ttcacgttat gtttaaaaat ctgtgctggg ctgtgtgggt tgcgtctgtg atcccgagcc  
28981 tttcgaggcg caagcgcggt ggatcacctg aggtcaggag ttgcagacaa gctctggcaa  
29041 catggtgaaa cccgctctct actaaaaata caaaacttag ccggcgcatg tgacatgcac  
29101 ctgtgattcc agctactcgg gaggtcgagg caggagaatc gcttgaaccc agdgagcgag  
29161 ggttgcaagt agctgagatc gtggcattgt actccagctt gggcaacaaq agtgaactc  
29221 catctaaaa ataaaaaaga aaaaagaaaa ataatacaag aaatagccg ggcgtgtgtg  
29281 cagcgccctg tagtccctcc cagctactca ggaagctgac cactctagcc tgggaaacaa  
29341 tgggaggtgg aggttgcaat gagctgagat cgtgccattg cactctagcc tgggaaacaa  
29401 gagcaaaact cagctctaaa aataaatagc ttgaaccggg caggcgaggg ttgcagtgg  
29461 ctgagattgc accactctat tccagctcgg gtgatagagg aagaacttat ctctaaaaa  
29521 ataaaaaat aactctttag gatggcaatg aalttaagga ctaaactagg gagaactcag  
29581 tttttttttt aaaaatggagt ctgtgtctgt cggccagact aggggtcagt gggcccatc  
29641 tcggctcaact gaaactccca ccttcagggt tcaagggatt ctgtccctcc agccctccaa  
29701 gtacgtggga ttacaggcac ccgccacact gectggctaa tttttgtatt tttagtagg  
29761 atgggggttt accatgttgg ctatggttgg caggctgggt ctgtacatcc ttgactggag  
29821 tgatctgctt gctctggcct cccaaagtcg tgggattaca ggcactgagcc actgtaccca  
29881 gccactctga catattttat ttattttatt attttttt tttttgaggt ggaactctac  
29941 tctgtccccc aggtcggaat gcagtggcac aatctcggtc cactgaaccc tccgctccc  
30001 ggggttcaag cgattctctc cctccagctt cccagtagc ggcataatga ggaacaccc  
30061 actatcccg actaatTTTT gtatttttca gtatagagga ggtcttccaa taattggccag  
30121 gctgggtctg aactcctgac gtacgttgat cctccacact cagctcccca aatgttgact  
30181 attaaagctg taagccagcg ggcctgggtg acatctttta ataactcgtc ttccattctc  
30241 ggtatattgt atatgtctcc atttacttag gtcttatttc atactctcta tcttggagct  
30301 atcatctttt ttcatacagg ttttgcaact ttcttgtgag gtttatctct taactgttca  
30361 ttgattttgt tgtgaattgg gaactctttt tccaccaagt atattttcta tcttgtact  
30421 ttgatataca ggaagataaa ctaattttta tctgagtttt attatctatg aaaaagataa  
30481 agtagaacta cctcagtaaaa ggtttccata atcaaaatag tatggggtaa acaagactaa  
30541 acagatgtgt tcaactgctg acttatcaat gcttgtgata attttttttt ttttttttga  
30601 ggcagagttt tgctctgtag cccagctcgg agtgcagtggt gctggatctg gctcactgcg  
30661 acctccactc cccgggttca agtgattctc ctgctcagc ctcccagatg gctgggagga  
30721 tggcatgcac caccacatct ggcataattt tgaattttta atagagacag ggttttccaa  
30781 tgttgactag gctggttcca gaactgttga cctcagtgga tctgctgccc tcaagctccc  
30841 aaagtctggt gattacaggt gtgagccacc accaccaggg aattgaaagc gtattattca  
30901 tgaagaaatg ggtagatttt aatgaacaa accctcttgg tgggcagatt cctaagctcc  
30961 agggccctcac acaaaagggg cagtgggctt ggagatccca gcttcaagct ccagagggac  
31021 tgctccctca gggccacccc agcccacttt tgatcccaa gttttgatca tcaagaatcc  
31081 caaagaagggc acaggaattt tcttcttcta cctgccatg aaaccttttg ccatagacga  
31141 tcttgaataca tactttggga aactgcatcc aaagacctt ctagtttcaa atctgtggt  
31201 ccagggtgtc ccaactgaacc ttacctgatg cccaaactcc caccattcca cctccaa  
31261 gaacacqaa gatgacctg tgccaaaatg aaagctttaa tgaattgtac tctagacag  
31321 tcaagtctga gcttctgcca gctccactg tccagctct ctgctgtgag cagcagggga  
31381 gctagtgtct gagggttagg gatctggagt ctaaaagaga gagccagga caaaggagga  
31441 caggagaacc cactggggg ctgggctccc ggaagtgtgt ggtgggggt ggtgtgggg  
31501 agggccggcc ctcttcaggc attctagca aagccacagg gggctccagg ggtcagagg  
31561 tccccatgg cacagggtgg gtgcgtttat gcttgcgcaa gtgctggga ctcagaaagg  
31621 ccttgggaga atgggggag gtgtaggggg gcaactgagc gtgagtcggg ctgtgtttg  
31681 gcagccggag ccggtcagag aagctcttgc cgaactRgt caggggaa gctgtgact  
31741 ccgggtgtga gcgctgtgc cgacgcagca gcgtcatgt ggagaaggtc tcttctgact  
31801 ctggcgacaa aactggggg ggcctctcgt cagctctcc acccccgcct gctgctccc  
31861 ctctcaaggag accagggagc tcccgagac cagcatctt gcatctcaa tgctccagg  
31921 tcatgccacc cactgccac tgtgtggcca tcacaccac tgactcccca gtcacaccg  
31981 gcagccctgc tggagggttc cccagccccc cccctgctRc cggggcggt gaaactctac  
32041 ctgccacgc cacaaggcag gccaaacccc ccaacagctc cgtgacagg gacacaccg  
32101 cggctctact gcttctgatg gtcgcctcgt gcttctcag gctgcagacg acccaaaagg  
32161 atttcccaaa tgcgttacag tggaaagggg gctcccccga gtgcacccgg gctctgtctg  
32221 tgaggtgtgc acgctggcg aaggtcgccc cgcactctc acagcggaag ggcctgact  
32281 cagagtgac cagcgctgc cgtctgaggt ccaagaaYgc cagcaagctg tgcgtcagc  
32341 cagggaactt gaacggcgcg tgcgctgtg gcacacgcc gtgcattgccc aggtccggcg  
32401 gtcgcgggaa gtccctgcgg caactctcgc agtggtatgg ctctcaccct tcatggggcg

32461	gctgggtggcg	acgggaagctc	gaggggtcgg	agaacatgcg	gccgcagcgc	gggcacagga
32521	agggctctct	ccccgagtg	gtgcgctggt	ggctctggta	ggacttgagc	ggacttgagc
32581	ccttgccgca	ggccggggcag	cgctagggct	tctgtgccgc	gtggatgcgc	tgtgtgccgc
32641	cggtgagga	tgagcyggag	aaqctctctc	cgactctcga	gcagaggag	ggcgctctcg
32701	cggtgagga	tgagcyggag	tggtggaggac	ttggcatgaa	ggcgacagac	ccataactgt
32761	acccactgc	ctgtctgggc	tgtaactttag	gggctcccc	aacgtctgtg	gggctctagg
32821	cttaactccc	tcaagagcac	atggctgcac	ccccagggaa	gaagccttca	gtgtgctgtg
32881	gtgtctctat	tccaaagacc	tgctcttgca	cattaaagac	caagatatgt	ggcgggcgcg
32941	gtggctcaag	cctgtaattc	cagcaatttg	ggaggcgag	gtcaggagat	cggagaccat
33001	cgctcaaca	cgtgaaacc	ccgtctctat	taaaactata	aaaaattagc	cagcgcttgt
33061	gggtgggtga	gttagtcccr	gctactcggg	aggctgagcg	hggagaaatg	cgtgaaccgc
33121	ggagggcgag	ctgtcgagtg	ggcgagatct	gtccactgca	ctccagcctg	cgaggcgag
33181	cgagactcgc	tctcaaaaaa	aaaaaaaaaa	aaagaaaaag	aaaaaaaaaa	aagaccacga
33241	cattggcagg	cgcggtggct	caegcctgta	atccccagac	tttgggagcg	cgaggcgggc
33301	ggatccactg	aggtaaggag	ttcaagacca	gtgtgaccaa	aacggagaaa	ccccctctct
33361	actaaaaata	caaaattagc	cgggcatggg	ggcgcctgcc	tgtaattccc	gctactcgga
33421	aggctgagcg	aggagaatcg	cttgaaccgc	ggaggcgag	gttgccgtga	cccgagatcg
33481	cgccattgca	ctctagcctg	ggcaaaaaag	tcaaaactcc	gtctcaaaaa	acaaacacaa
33541	aaaaacaaaa	caaaacaaaa	cagaaaaaca	agatacgtgt	ccccgcctct	ttttttctgt
33601	ttccccagc	tggaatgcag	tggcctgacc	atagctcact	gcagcctcga	ctccccagc
33661	tcaggcgctat	ctccccactt	atctctccaa	gtaccgggga	ctagaagtgt	acattcccc
33721	gctcggtgta	tttttttatt	tttataagag	cgaggcttgc	tgtgtgtgcc	aggtgtgtgt
33781	tgaactctgt	ggctcaagca	atctctctgc	ctcagctctc	caaagtgtgt	gaattatagg
33841	cgtagactat	tgtgcccagc	ctagaaacat	gtcaltaatg	taagagtgta	ccagaaagaa
33901	aaaaaaaatg	acctagacaa	accaggcccc	actcaacact	cctgtgtctc	acaaaagacc
33961	ctcagaactg	cccaactcca	aaccggcccc	cctttccagc	gtggcgaagc	cgacggccaa
34021	tctgaccaca	tcccatcttc	agagatcaac	ctcaaggctg	tgtccacctc	tgcccaatca
34081	ggggccacaa	tttctccacc	atgcttagcc	ctcccccttg	gatctgcat	cgaggcgag
34141	ccattggctc	actttaccct	gagactcaaa	ccccaggcccc	attggctcga	gcacgcgtgt
34201	cgccctgcgc	aggaaggcgc	gctccccgc	aaggcRccct	caccctctat	ggtgtgcgag
34261	gtccttgagc	tccgcatagg	ctttRccgca	acgctcacag	ctgtaggcgct	cgaggccagc
34321	gtgaatagcc	cggtgcttgc	ggaacactga	agggtcacga	aaqctcttgc	gcagctgtgc
34381	gcaggcgtaa	ggccgctcgc	ctgtgtggcc	acgctgggtg	atctttagct	tggagagcgc
34441	gccatagccc	ttcgggcagt	gcgcacagcg	gaaggcgagt	tccccagcgt	cgagggccag
34501	gtgcacgcgc	aggcacacgg	gctgcatgaa	cgccgcggcc	cactcggggc	aggaaaaggg
34561	cttctcccc	gtgtggctgc	gcccgtggct	gcgcagctcg	ggtgcgctgt	tgtaggccctt
34621	ggggcatagc	ggacacgcgt	aggggcttag	cttggcgccg	gagcctgaca	ctctctcccc
34681	actggctctc	tctgctctag	cttctgtctc	tggctttggc	tcaacctggg	ccacctcttc
34741	agagcagctc	gcggcccatc	gtgtggcagc	gtggcgcgct	gcctctggcg	cgtttggaaa
34801	tgtcttggta	caggacaSgc	actgttagcg	gcggcccgag	cgcttgtgaa	cgggggctgtg
34861	ggacgcggcc	tctgcagcct	ccacttccat	ggccttggtg	aacgggggtt	ctctgcaaga
34921	gaagcaaatg	tagacaaaag	ccactaacct	tgcacctccc	tgaagaccct	acaggaaccc
34981	ctatctctat	tgcattttct	acctgggaac	ccacacatcc	tctctgccc	gcatctctgt
35041	ctctgacatc	ctgcgttctg	tctctccctg	atctgtctat	tgaagaagag	cttctgacca
35101	agtgctgcaga	gagccactaa	gaaaggaggc	tgagggtcac	aaaagattca	ctcacacctc
35161	ccccccScc	cccaacgggc	ttttccaaac	actgtggcat	tcccagagga	ccagggttca
35221	tctgtctcac	catcttctct	cttcagctca	gtgtccaaag	actatgccag	tataaaggat
35281	gtaccacagc	gtgggcctg	cctgaaggct	tccaaagcgc	cagaaaaggga	agatcctggg
35341	ccagaaaagg	gctgagccaa	tgggctaaat	ctgtgtagctg	gcactgtctg	gaagtgcagc
35401	gtccccagca	tttgtgtttt	cttctctctc	ctggaaggtt	agtcctcaga	gacagcaact
35461	gttcaacagc	aatctgtgcc	ttgcacagct	gtacggccct	ccggcccgaa	ctggaaatctg
35521	tccactctct	gactgtgaa	cttgttggcc	tgttccacac	actctgtgaa	tctgaaataca
35581	ctcaaggcag	cctgagccat	tgctagcagc	tggaaagcct	ttcttgatct	ataactgatg
35641	tatctgatct	aaactggcct	cctgggatac	cagctcttagc	tgaagatccct	actttctgtg
35701	ccagaaccca	gacgccctcc	accagctgc	tcttgggagt	catgtgtggg	aggaacacgg
35761	atlaattggct	gtattagtct	taacaccagc	tcatctcccc	tgggggtaga	agggaagagg
35821	atlatggcct	atccacttaa	ggagtgtcta	gcagctgctg	ctgggggaag	gggtctgaga
35881	agtgggggct	gcaggagacc	agggtgtgcc	agagggctag	ggggctacgt	gttacttgcg
35941	gccctgtgga	ttactatgag	acctcagta	aatgagtgtt	gtttataaga	ctatttctcg
35981	ccgctgtggc	gtggtggctc	acgctgttaa	tccagcactc	tgggagggct	cgctctctact
36061	gtcacaggag	tcaggagacc	gagaccatcc	tggtcaaac	agtgaacccc	gctactctgct
36121	gaaatacaaa	aaaaattagc	cgggcggtgt	ggcgggcgcc	tgtagtccca	gctactcgctg
36181	aggctgagcg	aggagaatgg	tgtgaacctg	ggaggcgag	cttgcagtg	gccaagatcg
36241	gtccaccgca	ctccagcctg	ggcgacagag	caagactccg	tctcaaaaaa	aaaaaaataa

36301	aaagactatt	tctcatctg	gacacgttag	gggttggtg	cttctaaagt	atgacactgg
36361	gggtcaggag	gaggaactg	agctgagcca	cgagggtgct	aggggaaata	cgagactcca
36421	ggggccctta	tgccaggaaa	ggcgaggaga	agcagcttac	ctgggtgacc	cgaggaggac
36481	acagaggccc	ctgacgcgtg	gcaaggggcca	catctctaaa	gtcagcacc	cttgcctcga
36541	aatccccggc	tgaggccctg	acgtctgcc	ccctcccaaa	ccccatgcag	ccggtctctt
36601	cccaaaagta	atgaatgatg	tttctgttcc	ctgtctcaag	aaacttccat	ggcttccacc
36661	acacacttga	aatccccagg	tcttctctgc	ttgttccaaa	tctgacaaat	cccttctacc
36721	agggcccaat	ccaccaggag	gatctctca	catctccctt	ttgggtttgg	cccacatccc
36781	acaaaactcc	atcttatctt	cacttggtta	tttcaacttt	cccactaaat	cctagggctt
36841	agccaggtgt	gggtgcctgc	gccttaagtc	ccagctacgt	ggggaggtga	ggggagagga
36901	tctgtgaac	ccagggggtt	gaggtgcag	ggaactatga	tgttaccact	gcgttccagc
36961	ctggggcaac	tggtgagacc	ccatctctac	taaaaaataa	aattagcaga	tgtggtggca
37021	tgtgcctcta	gtactgccta	cttgagaggc	tggtgtggga	ggactgtctt	agcccaggag
37081	ttcaaggctg	taccactgca	ctccagcctg	ggagacagg	caagacactg	tctccaaagt
37141	tcaatcaat	aataatcaat	cttggggctt	gaagataagt	taaaaggact	gaattctaac
37201	ctttctgatg	acttgaattc	tctctacagt	ttccaaggga	tctctcccta	tcttctgatg
37261	aggtactcac	tacctcttcc	agacggtttc	tggtgagctc	ggctgataat	gttctctctt
37321	aataaaaaat	atagcttggt	ccaggttccg	tggtctcac	tacactcttt	ggagacaccc
37381	cggttgtagt	cgcttgagcc	caggagttca	agacaaaggc	gggtgatact	ttgagccag
37441	gagttccaag	ccagcctggg	cgatatagca	aaatcccat	ctatacaaaa	atacaaaaaa
37501	tagccaggca	tggtggagca	tgctgtact	cccgactact	ccagaaggct	gaggttagag
37561	aaatgcttga	ctctaaggag	actgaggttg	cgatgaagca	agatgggtcc	actgcactcc
37621	agcctggca	cgaggtgaga	ccccatctca	aaataataa	ataataataa	tgtataaaaa
37681	gatagtctac	atttactgag	cactcgccaa	ataccaggca	gatctctaaa	ctccttatgt
37741	gtattagctt	agttaccctt	catggcaacc	ccatgaggaa	agttcttata	ttcctatctt
37801	acagataagg	gaaccaaggt	ccagagaaat	ggttcagtat	tttgttaagt	gcccgattcc
37861	tgaagccaaa	ctgtctggct	tcagatcttg	ctctcatcac	ttcccgactg	atgtgacctg
37921	gtgtaagtga	ctgcagtctt	tagaacctca	gttttcta	gtcaaaaatg	gagataaagt
37981	Yagttactac	ctgacagagt	tggtgaggga	atgaatgagt	caaaaataat	tactgtctct
38041	aattatacaa	cgctcttctt	agagccagac	acattgctgg	ggtgtgcgtg	tatttataat
38101	taattgtctta	taScttYctg	agatagggat	tcttgcctct	actttacaga	ggaataaact
38161	gaggttccaa	gagttaaagt	acttggccat	gggtccacaa	tggttaacct	aaacagctgt
38221	gacttcaact	caggtgttta	atttgcctta	agttgctggg	gtcttgcctc	gtgtgtctgg
38281	gcctcttaag	cttgcctgct	gcctccgcca	gcctccagct	gacccagacc	ctgagctctg
38341	gtcataactg	tgcttctctt	catcttgca	gaactgcctt	gagacccctg	ccggccacct
38401	ttatgtctct	gcttccctct	cagagggctt	gacccagtgg	tcttgcctct	gggaccttct
38461	actgcctctt	gccagctctg	Kgtctcagcc	ttcctatctg	tgagttagac	accaggttag
38521	tggtggggaa	atccctctct	ccatggcact	tccagggga	aaaggttagg	gagtgccagg
38581	ttgtgtctag	catgcgcccc	gctacacaaa	gaggtgaggt	aggttaggtc	gttctctaac
38641	atccccacct	taaaaaaaaa	aaaaaaaaaa	atatataat	atatataat	atatataat
38701	tttgatagtg	agttcttact	tggtgccag	gctggagctg	agtgccgcca	tttgggttca
38761	ctgcaacctc	cgctctctga	gtagctggga	ctacagggca	ctgccacac	acctgtccag
38821	ttttttgtat	tttttagtag	gacgggtgtt	cactgtgtta	gccaggttag	cttcaatctc
38881	ctgacctcgt	gactccgacc	ctcgccctc	ccaaagtgtc	gggatcacg	gcgagagcca
38941	ccaMgcccc	actttttttt	ttttttttat	taaaaggtct	gaggttaggc	tcaggaatct
39001	gtatttttaa	tacactctgg	atcattccag	ccaatacttt	tgtttgtttg	gttttgaac
39061	aaagtctcac	tctgtcgtct	aggtctggat	gcagtgggtg	agtcactgct	tactgcagcc
39121	tgactctccc	agactcaagg	aatcctccca	cttcaagctc	ccaagtgtgt	gggactagc
39181	gcgtgcacca	tgatgctgtg	ctaattttta	ttttattttt	agtagagagc	aaagtctgtc
39241	atgttgcctg	gcttggtgta	gaactcctgg	gctcaagtga	tcttccctMc	gcagcctccc
39301	aaagtctcgt	gattacagcc	gtgagccacc	gtgcttcccc	aaaccaatct	taagagccaa
39361	acacacatct	tttaggtctc	acgagctctc	aggagaggag	catttataag	gttctactca
39421	ctctcttttc	agatatttag	attaaggtcc	ccacaaagca	aaaactgtac	acaaaggac
39481	acagctgttc	aaggagccag	actogaaccc	aaagtctcaat	tctctccccc	aggttgaact
39541	atgagactct	ccactgtctc	caggaaaaag	accaataatc	tttccatggc	cagctagacc
39601	ccaaacctct	taaatctctg	ctacctgggc	agatcacttg	agggcagagc	tttgagacca
39661	gcctggccaa	catggtgaaa	ccccatctct	actaaaaata	cacacacaca	cacacacaca
39721	cacacacaca	cacacacctg	cttacctcac	ctccactccc	tctccctggc	ccactgggct
39781	ctaccacacg	agggcctctt	tcttctctcc	aaagagctaa	attccttccc	acctcagggc
39841	agtggaacta	cgagtctctt	ctgtctgagc	cactcttctc	ccagactctt	tgcttaagct
39901	tctttttttg	tgttattttg	atctcagctc	ccagtcactc	ccctaaaaat	agctttctctg
39961	accacacttc	cttttctctc	ccccctttaa	tatttccaaat	tttttctctt	tttaaccact
40021	caaggagcac	tgaatgacta	ctttctctaa	tgtctatttt	acccttgata	atcatctctc
40081	atctactctt	tatttatttt	attatttttt	gacggaatct	catgatatac	tctcaagcag

40141 ttctctctgcc tcagcctcct aagtagctgg gactacaggt gcccgccacc acgcccgggt  
40201 attttttttt ttcttttttt ttgagacgg agtctcaact cgtcacccag ctggagatgc  
40261 agtgggcaaa tccgtggctca ctgcaagctc gccctcccggt gtctatgcca ttctctctgcc  
40321 ttacgccctc gagtagctgg gactacaggt gcccgccacc acccgccggt aattttttgt  
40381 attttttagt gagtagctgg ctactctgtg tagccagagt aggtctcaatc tccgtacctc  
40441 gtgatccacc tgcctcggcc tccagagtg ctggagtac agggctggagc cactgacccc  
40501 gggcccaatcc cggtcaattt ttgtattttt agtagagagt aggttttcaac atgtgtggcca  
40561 ggaagtgtctc gacttatgta cctcgtgato gcccgctccc aggtctccag agtgcgtgga  
40621 ttacagaggtg gagccacccg gcccgccctt ttttttgaga cggagtctta ctctgtcccc  
40681 caggctggag tgcattggca caatatctgc tcaactgcaac ctgcgctcag gggttcaag  
40741 cggttctctc gctcagcttt ccgagtaggc tgggaattata gggcgccgcc actacactctg  
40801 gctcattttt gctatttttag tagagagagg atttcaccat gtgtgcacag cctggcttga  
40861 actcctgacc tcaagtgtac caccaccctt ggcctcccaa agtgccttga ttacagggat  
40921 gaggccaccg acccagccct ctttactttt taataaattgt tttattttt atttatat  
40981 ttatttttga gacagagttt caactctgtt gccaaaggct gagtgcgaat gccacctctc  
41041 tgcctaccca aaactccgcc tcccggttcc acgagatttt gtgtgcacag cctcccgagt  
41101 agctggaaatt acaggctacc accaccacgc ctggttaatt ttttgtatt ttatagtaga  
41161 ttggagtttca caactgtggc caggctggct tgcgaactcc gattccaagt gatccactgc  
41221 ctccgctctc cagagtctgt ggattacagg catgagccac cgtgctcgtg ttattttttt  
41281 ttattttatt attatttgtt attattattt gagacagagt cctccctctg tggccaggat  
41341 ggaagtcaat ggtgtgatgt cagtccaact cgacctctgc ctccYgggtt caagcaatc  
41401 ttctgtctca gctctccaaq tatttgggat tacaggtgcc tggccaccaca gctcctcaat  
41461 ttttttgtatt tttagtagag atggccatgt ggtgtaggct ggtctggaaac tccgtacctc  
41521 aggttagtcca ccacacttgg cctcccaaag tgcgtggatt acaggtctga gccacctatg  
41581 accgctcatt tatttctatt ttattttatt attttctttg agaaaagtct ttgtgtccc  
41641 aggtgtaggt ccagtggctg catctcagct cactgcagcc tccacctccc gattccaagt  
41701 gattctccag cctcagctcc ccgagttagt gggactacag gcgaagaacca tcaacctctg  
41761 caaacttttg tatttttagt agagacaggg ttccaccaca ttggccacag tgcctccga  
41821 cctctgacct caagtgtacc gaccgctcca gctcccaaaa gtgctggagt tacaggcgtg  
41881 agccacagct cctggcctttt atttttattt agagattggg tctcactctg tcaacctctg  
41941 ttgcaagtgc tcaatcatag ctcaactcag cctcaaaactc ctgcaactca gcaactctcc  
42001 tgaagtacta gacactatagg caccaccacc cacaactcgc taactattata aaaaattttt  
42061 gatataagat atggaggtct cactacgctg ccgacagtggt tctcaagaac cctgctctc  
42121 agagttagct ccaactWtggc ttctaagaat gctgtgatta caggcatgag ccaactgac  
42181 cagcMccccct cctctcataa ttcttttatg gtgtcactct gacataactc ctgtcaagct  
42241 tccactgggc aaggtatcat tctaagcatt ttgtgcataa tactcaacta ctcaagccaa  
42301 ctgcacagct gcttagcagt tcatatgagt tgaattgttt gtgtccctgt cctttcatat  
42361 attgaaattc taactctcat tgtgactgta ttggagaca aggcctttat ggaagtattt  
42421 aaggctgaac gacRtcaataa ggttatggcc ctgttccagt aggattatgct tctttatgct  
42481 aagtgacacc acaagagaaac aagettacta gtccaggtcc cagccagtggt tcaaatccca  
42541 aacacctgct ctctgagccc tgactattgc ttgtctagca ttacacatct tatgttttgg  
42601 ctgttaattc ttcatcacca gaccagagt ccaggctggc aaaggcgtag aaaaagcgtg  
42661 atctgctccc tctacaccca gaacctctgt tggtagccca aaacaaatgg aaagaaccaa  
42721 cctcagatct ccaattgaaac caggctctga cctctcgtct cctccacagt agggagtgtg  
42781 gagaatgtcc accaataact gtttattaa taattctctc cacttttcca actccacaag  
42841 gctcaactat tgccccata tcccacagtg ggtcctctgt cgtggcgatg acctcgtgct  
42901 ttgctcactc actggtctag gaaggggatg aagtttgtct gtctcactag gcaagcaaa  
42961 attcaatttc caaaaactct aggtctaggac cctggggcag cctgtccactg cgtccagag  
43021 cactcactat tctccactct taacaagca gtccaccac aggcctccag gccagagagt  
43081 acgtagttag cggatgaacc ctgcagaagt gagcgctaat catatcact cctctcacc  
43141 gggccagctt acaactgat cccaactact ttctccagcc tcaagctccag ccgacagagt  
43201 cctctccccc cagcgaccca acctccagcc acaattctca cctctccacc ctctagatg  
43261 gtttctctgg ctgggggagc ccttgctcgt ctctccacc gggaaagccc ctttataccc  
43321 ttcaagaacc ccttcccta tgcacctaga gaataactt taaggcagag agggtaaagg  
43381 ccccaagccc ctctggacct cagtctctcc atcagtaaaa tggggctcct taagtttgtg  
43441 ggaagtttga gggcggtggg gctcagaaag ccgcgcgag ccgcagagcc agtcgaag  
43501 cggagctctc cgtcctcgc accgctctgt acgcggcgcc cccgcagacc aggcagcccc  
43561 ttgagggtag gaccocggtt ataagctcca gaataagctg gaataagctg cgtcgagga  
43621 aggactgaac gaagatgag gagaagaaca aatgaatgga gcgaagccgc attagggag  
43681 aaaaagctga aagatggcgt tctctaggac cgcggagagt gtggccggag cctagcaggg  
43741 gaggggggcc cggcgccgtg ggaatcttga cgcggcgcca cgcagccagg accaaacccc  
43801 aaaaactggc gggctctag acccccggga ccacacccgc ccaccccgag agcccgccgt  
43861 gccgagcctg ggtgtccgca gcccaagaacc cggcgagagt cggcggggt cttagaccgt  
43921 ctggggccgc aattgcgccc gggcgctcca cagcccgccc cggcccgctg cgtcgcgctg

43981 ggcccaaccg gccgcgccgc ggcgcccgca gggaaactga ggccaccctg gatcgtggcc  
44041 ggcgtggagg gccgcggggc ggcgcggcgt tccgcgcgac tgatgtgaac  
44101 tgcgtgact ccgtggcgct ggcgtgcagc ccgcggcccg gacctcgcca  
44161 aacaggggcc ggcgctagga cccagcgagg ggcggggagg ggcgcgagc  
44221 gcggaagtag gggatctccc tcaactagga aggaaggagg agttcccgag gaacctccag  
44281 cctatggcgg cagagaagca tcttgcaaga ggtctctgtg ttgtctgagg caaggggagc  
44341 ccaggcgagg tgacgggata cggccgcctt gttgtagtct gggggcaggc caccggcaat  
44401 gctctcaaga accagaaggt gggaggagaa aaaggccatct ctcaagctctc gcaaaacctt  
44461 agtccaatta tattccacaa catttccctt aatagcaagg agtgggttct acgaagtctt  
44521 cctccatctg tctgggaaaa gctcctcaga gaagcccgag aagagtgta ccaaaagtacY  
44581 gttgcggagg ctgtagaact ttttctcccc ttctgcgccc tattcaaat gcaaccggaac  
44641 gattctagct gctcttgccc gaagtgtggc tagcWtagc tgcgaagagc ccggtttgtg  
44701 acctacgtgg gcggagtcag ccgtgcagac ctagaactta gcgcgggaag tgtgtctggc  
44761 taaaacttgc cctcggaggc gggacaggca gtgctccccg agcggaagtt tcgcggggaa  
44821 gcttttgacac ctaaggacac ttctgttttc ctagtacaaa taggaagtgt ccgtagagct  
44881 gggagtagac tctggctcca tgccagctgc gccctctttt ctctactctc ttctctctcc  
44941 cccctctctc ttctctctt tcccttccct cccacctccc gggaaccctcg gctgagtgtg  
45001 cgtgtgtggg agcgcgagag ccccccagca gccacctctt gggcgccgag gctgcagtta  
45061 gggtaagggt ccaggtgcgc aggcgcgcctc ttgtctcgcg tcccacagct acgcgcggcc  
45121 ggcgggaagg ccaggggggc gccggtgcga ggtaggcgct ctgaagagaa atttgcagac  
45181 gagatgagtt aggtcagagg acctcattgg tggttgacta ggggagggga ctttgtatca  
45241 aaagggcgct gtgaggtgag aattccgggg gaaggttcaq tgaagggttt tacygtcgga  
45301 gacagctgta ctgggggtccc atgggaagaa gactcagagt tggagaagga caagacagat  
45361 tagggggcat gagagagcaa gggctttggg atgaagaccg cctcagtcag ccggtttgtg  
45421 gttactgtga ctagaataac aggcattgccc gttgtttctg aggatgtagt aacacaatga  
45481 atatagttgc cactgctgct ggcacatggt aataataac ttctcagatg tctgtctcat  
45541 catcagttat aagagaggag gcagagaagaa aaatgagaa agactggtga gggtaagatc  
45601 aggaagcgat ggagagaaaa ggtaggcaat gaagtgcact tacaacctgg tattgtggtt  
45661 attcccaaga tggagaatag tttagtttca aggaagtag ataaagaag tcgctgaagag  
45721 agtcttttgg gtttttttKt ttttttgaga tggagtctcg ctctgtcacc caggctggag  
45781 tgctgtggtt ggaatctcagc tcaactgcaac ctccactccc cagggtcagg cagtcttcca  
45841 gctcagcgct cctgagtagc tagaattaca ggtgccacc accaaccggc gctaattttt  
45901 gtatttttag tagagacagg gtttaccact gttgtccagg ctggtctga actcctgaac  
45961 tccagggtac caccagctcc ggcctcccaa agtgcgggga ttcacggctg gaccacacca  
46021 acctgcacca gttaagagag tcttaactct cttaactctc ttgtcacaag gaaaagagac  
46081 cttgtgacac tgaatatgact cggggtggtg ggggaacaa ccaagccctt cctgaaagga  
46141 ggcctctaac ctctcctctc aggtctctcag ctattaactg gaggaaacag ctgcttttct  
46201 agtgctcttc agctactctg tttagctgag agatgaagt ggaagatttg gactctctct  
46261 attgaaaggc ctgagaaggg tttgtgtgct ctttaagat gtcacagaaa attttgtgtt  
46321 caggattgta gggagcagat tctactggtt cttaaaggac agtaaatgctt tttagtctg  
46381 gtctgaagaa cataacaggt ctgtgatcag aagtagggtt catctctctc aactttaaYt  
46441 tctcttagta taccgttagg gatgacttaa gctcagggga gctctctat ttgtggaagc  
46501 tgtgcacagg gaagccttaa atgatgggtc ctgcagattg gatctagtat aaattaggtc  
46561 cttggcgctg gatgcttggg gaacctctca gtgacctcag gtaacctgtg tgctctgag  
46621 gccaaaggag gaagttaatt caggccctcc ttttgaccac tgccccctct tccataggct  
46681 tggccctccc accagaggaa ggtgctgcca ggtgctgcca cctctgtaac cctcagggtt  
46741 ctgtctagctt gcccatgga ggcacacacc cctcactcca gctgtctcga ctgtcagcgc  
46801 cactttccca gctccccaga gctctctogg ccacgagaa gctgtccatc ctctcccaac  
46861 caggacagtg agyaggtgta cagcatccct cgcgtctcga ccacagagac gttggtggcg  
46921 gctacagctc accccgggag cctggttaac catcgtcgga cccacagagc ttggctttt  
46981 cccgtgacca cctgtggcaa ggaactctcc aatccatgg ctctcaagag gcccaagga  
47041 acacatagct ctgaggcccg ccgcaggcac agggcccccac gctcctcagg ccatatgag  
47101 cactccagg gtgagacggt gtcactgac tctgtggggc aaaggtcttg gctgaaat  
47161 ggcgtgggaaa accagacaaa acatacagaa ctgcccacca gacaaagaga aggttgagga  
47221 cccaggccag cttcgggtac ggtgggaagt ggtgagagac cctccactga gctctgagga  
47281 agccaccagc gctctgagga agccagcagc ctcttagaca acttggaaca gtaactggct  
47341 gccctctctc tccccatccc agccagcagc gaagccaccc agactccctc gactgarRag  
47401 gaatcagtag tgaacttccc agggggccag aagactcata agcagccggc gaggctcacc  
47461 gaggcggcgt acaaatgtag tcaagtgtggc taacctgtgt ccaactgtgt gccatgccac  
47521 aaccacccgc agagccacac gctgggcatc cgaactgtat cccagatctg gccatgccac  
47581 tctaactcca tggctctgaa cttcggctct cctcgggagc ccccgaggag caccagctcc  
47641 gttcccccct gcccccgtgt cttcgggagc gaggRgaag ggaagccac caccatggg  
47701 catgaggggt aaaggcagga gccacgctgg ccaagtgcac agatgctgaa tggctgtcgg

47821 gagctcagca cctctgggga gctggaggag agtggcctgg aggaataccg gcccttccgc  
47881 tgtggggagt gtggccgtac ttaccgccat gctgggagcg tcactcaacca tcgaaagagc  
47941 ctcacagacag gtgtctaccc ctctcactc tgtctcaag agctgttcaa tcgaagctgc  
48001 ctcacaaaacc atgtgcccgc tcatccacag cccaggcaag gagtgtggga aagtgggcag  
48061 ccatcagctcc caccagctcc ctactcagcg gctgagacca cccacaagag ggaagagcag  
48121 cccaccacca cctctggacca tgggctctat agtgcagtgt cttgtgtctg tgtctaccgc  
48181 caccggggga gccctgtgaa catctgccac agcatcgga ctggaggtga ccagtgctca  
48241 ctctgtcccc gcaagtaccg caatctcat gcccctgcga accagctgca ggtacattgc  
48301 aagctgtctc gcaagtagtc agacatcggt gctgagggtg cccccagcca acacaagta  
48361 gaactccctg tagaccagtc gtaggcagag cgaccccgc cagacatca ggcactgtg  
48421 tgaacaactc aagaagagtc caccggacat accccagcag cagacaagac acgacacat  
48481 atctgtagca tctgtgggct ctctttgaa gacgtcgaga ccttgaagc ctatggcctg  
48541 actcatgggg caggggaaaa ggaataatag agaaccagaga ccaaatgtc acctctagg  
48601 gcccttggct gccgagactg tggaaagagc tatcgccact caggcagctc tataccac  
48661 aggcagaccc accagacagg agacttcaat tgtggggcct gtgcacaaga cctccacc  
48721 atggctgcca tgaagaacca ctctgccggc cacagtccgc ggcggagcag ccgctctagg  
48781 aagcgggctg gcggtgccag cgggtgggga gaagccaaac tctgtgcagc ggaagactgg  
48841 acccgggagc tagaagacca tgaaggcctg gactctccc aagaccttc aggggaagt  
48901 cctcatgggg ctgaaggcaa cctggaaagt gatgggact gtttgcagcg tgaactcgaa  
48961 ggggacaaat gtggccttga gagggatgag acccaattcc aggggtatga agagagcag  
49021 ggcactgggg aaggactgga aaggaagagt gccagtttac ttgacaactt ggaactccca  
49081 ggtgaggaag tctgtggcac tcaactctgc gatagctcca ctgggggtga tgaagacag  
49141 caccagacca ctggcccaacc caactcctct tcccactctg ccaatgtgct cactgtgctg  
49201 aagctggggc ccgtctacac atgctctgac tgtgggcatt tcttcccaca gctccactgc  
49261 ctctgtagac acaggccctg ccaaccacca ggcattctat agtctccctc ctgcccgaa  
49321 gattttgact ctctgctgct ctctccgagc cacttccaga accatagccc tggggagcgt  
49381 acctcagcac agcctttctt ctgctgcctc tgtggcatga tcttccctgg cgggcggtgc  
49441 tacaggcttc accgpgccca ggccacagc tctctggca tgaactgagg tccaagagag  
49501 gagggggaag aggaaggagt ggcagagcca gccctgcac gcactccacc actgcagctc  
49561 tcggaagcag agctgctgaa tcaagtgcag cgggaagtgg aagcctggga cactgacagc  
49621 tatgggacca tctgtggctg ctgtgtgtcag acctacgat cctggggagc cctggagcgt  
49681 accaccacaa gtacagattc tgggactact gcagacaagg ctcccagccc ctctggagct  
49741 gcaggtgatg ccatggagat ggtcgtggaag agtgtcttgg aggacatagt gaattctgtc  
49801 tcttggagag gtggagatgc caagtctcaa gaggagcag cccgccctt cgggagacac  
49861 ctctgcatcc aggggtggga aagtttgtt gaggctcagc cccgccctt ccgctgcac  
49921 cagtgtggca agactatgc ccatgggggc agcctggtga accacgcgaa gctccaccag  
49981 actggagact tctctgccc tgtctgctcc gcctgctacc ccaactgtgc gccatccgt  
50041 aatcatctgc ggaaccaccc tctctgcaaa ggctctgagc cccaggttgg tcccatccca  
50101 gaggcagcag gtacagtgta gctgcaggtt gggcccatcc cagaaggagg ccgacaacag  
50161 cccagcacca tggcagagga ggggcggggg caagcagaag tcgagaaggga cagcaagaag  
50221 cttaaatgtg agcccttgga ggaagtggcc aggggtgaaag aagagtggtg gggagggacc  
50281 actgtgaagg gggagagat agagcccaag ctggagacYg ctggagaagg ctgcgagact  
50341 gaagccagct ctgagcgccc ctctcagctg gagggtgtgt ggcgactccta caagcagccc  
50401 ggcagcctca tcaaccaccc gcagagccac cagaccggcc actttggctg ctacggctgc  
50461 tccaagggct tctcaaacct catgtccctc aagaaccacc ggcgcatcca tgcagatccc  
50521 cgacggttcc gctgcagcga gtgtgggaaag gcttccgccc tggcggaagg cctggccagc  
50581 caccagcggg tcccatgga acggcgctgg ggtgggggca cccgaaaggg gactcgggaa  
50641 gatcgccctc tccgtgtggt gcagtgcggg cggaactatc gccaacggcg gacgctcctg  
50701 aaccaccRgc gcagccacga gacggggcag tacagctgccc ccacgtgccc caagacctac  
50761 tccaaccacca tggccctgaa ggaccacag aggtgtcaact cagagacttg cggcgacag  
50821 ctgtgagcgt ccaggcgccac agctgtgcgt tgcgccctct gtggcgcaag ctctccctgg  
50881 cggggattct tggagcgcca cctgcgggag catgaggaga cactgactgg tggagatgg  
50941 gggcaggagg gcttgatgg cacagcgccc agtgaggcga acctgactgg cagccagvga  
51001 ctagagaccc aattgggtgg tgtgagcca gtaccacact gtacggatgg tgagactagg  
51061 ccaggggagc gcagtcaag ccccatcagg ggggtagggt cagagagccc cagactcact  
51121 tcttgggtg cagggaaggg aggtgggtgg cggtaggtg ggggactagg ggggactagg  
51181 ggaRctggg tctctcagtt ctaactagg tcaaggagc cagagagcag cagagactag  
51241 agtctctgcc acgctggtga ctgcccagtc aatggaccta ctctgactca ctctgacag  
51301 tgggacacaa gagacaacag ctctcagctg cagccaggga gccaactcct ctctgacag  
51361 tgtggcagca cttaactgcca gtcaggcagc ctctgaacc acaaccacaa caagacagc  
51421 gcacactatt gctgctctg ctccaaggag tctttaaact ctgtggccac aagaagcacc  
51481 accacacacc acctagacgc ccagaccttt gcctgtcctt actgtggcaa agccttggag  
51541 tccaccagg aactggccag ccaactgcag gctcatgccc gggggcagac caggtgcca  
51601 gccagatgg aggagggcag agatcccaaa gccgggactg gggagagcca ggtggttctc



51661	cttggctcaag	ggaaagccca	ggaggcccca	tcagaaaccc	ccagaggccc	aggagagagt
51721	gtggagagag	cgaggggag	acaagcggtg	acgtccatgg	cggtcgagga	caagggcggt
51781	cccttcctgt	gcaccaggtg	cgggcgctcc	taccgccatg	ctggcagcct	gtcgaaacac
51841	cagaagggcc	acaccacag	gttgtaccgg	tgtccctctc	tgcccaaaat	cttccctaac
51901	ctgtctgtct	ttaaagacca	cagcaggacc	cacacggacc	caaagcgcca	ctgtctcgag
51961	atctgtggta	aggcctttcg	gacagctgcc	cggtctggag	gccacggcga	ggtccatgca
52021	ccccggggag	ggcctttcac	ctgcccccat	tgtccccgcc	acttccggcc	cggatcacgc
52081	ttcgtgcagc	accagcagca	gcaccaggag	gagtggacgg	cgctggcgctc	cggtaggggg
52141	catgaagggt	cccaggagga	ggtgggcaca	cagtggaggg	ggaggtccag	ccccaaagtc
52201	gggtggggag	caaggagtga	gaggagagag	ccccgggat	tctaagaggt	gggtgggggc
52261	tttgtctatg	ggtgagagaa	gtagcttgag	gatgtgtctg	gctgagcacc	cgcaagtcat
52321	gtataacaaa	taacaggggt	ggttgggcag	cacgtggggg	ctgtgttcagg	ccagggctctc
52381	tactctgggt	ccctccattac	actgtagcca	gaatggaatg	gtctttctgt	tcaggggagag
52441	gtcaactggg	acccctctgc	tgtgtgtgct	ggaacacctc	cgatgttcag	cagataaagt
52501	atgacttcca	gagaaaaaga	ggaagccatt	ggtttgtgtc	aggttccatt	ctttctctga
52561	cgaggccggg	tgccagggaa	caagggatgg	ggcattgggt	ccagcgcttc	ggtgtggaat
52621	tgccccagca	taactggtca	ctggttgcca	ccctactccc	tgctccctct	tcctctcgcc
52681	ttgtctctgc	tgtctctctc	cttggaacct	agactctlgt	tgctccctgt	tcagtgttgc
52741	tcctcatctt	tccttaacct	ttattcagcc	ctttttccct	ctgtgcccac	cgcccttttt
52801	aggatccaac	aaacacaccc	tttctactcg	cgacccctgc	cagctcttca	acacctttaa
52861	ctgtgaggat	gagtcacaga	taattgtttc	cttgaagtcc	agggccagct	cgaccaacaa
52921	cagtctatag	cccggtgcac	atccctgac	agagggatcc	tcggtgggga	atgccttcca
52981	cccagcactg	ctggaagccg	cRgctgccag	ggagtggggc	ggcgctgttc	ctcagcagca
53041	ctgtggctgt	cccttccaac	tcctctagta	gaggcggaac	ctatccatct	agtgccgacc
53101	gaggggtggg	ggcccttgaga	tggtggggcc	ttgacaggcc	ttgtcagagc	agaggggcaag
53161	tggtgagcat	ctgaaagctg	aaggaaatgg	tttaaggata	gaagatttct	cagatctcca
53221	agggatatga	gggaggagcc	agtttggccag	ggctgggaaa	ataataggga	ggcctagaaat
53281	ccctgttctc	atctgggccc	cgggggccag	gggaggggga	aaagctcgta	cggtctggag
53341	gggtgtcacg	ctgtgcccgg	tctgccccct	agttggtgac	ctctctctct	ttctccccca
53401	ggagccccag	tggcaccagt	gacgggcaga	gggacttgc	cattgcccc	tcacacccac
53461	cccacgaccc	caactcctgga	tccttcaacc	cagtgccctg	cagacatcag	cttctccctc
53521	tgaaactcaa	gtctccaaa	atcagaatct	ggggagggga	ggcgtgggga	ggaggggctt
53581	gatctccaca	ttttctcagg	agtagttcgg	gcatcccatc	atcttctctc	ctccccYtct
53641	gaagaggagc	cagatctggc	ttcttttcca	agggaggggt	gggggttctc	tcgcgtccag
53701	gtccttgaag	gacctctctc	ccccagcttc	atcaccgctc	ttcttctcag	gccacccctca
53761	gcagccagat	tgcaaacaca	gggagaggcg	gatgcagagc	cccacccggt	ggaaagtgtc
53821	ctgtggaagg	gagccttttg	ctacaaattg	taactatttt	tctaaagtct	attttgttaac
53881	aatttattta	agtttaaaaa	aaggaaaact	gctgcccccc	aaaaaaagaa	attttcaaaa
53941	caacgtggct	ggcgtgatgt	tatctgaaag	ggtaaaggag	gaggaaaagt	gagacgcctg
54001	cttgtgtagca	gagttgggtg	tgaggagtgtc	cacagacacc	cctgtcctgc	aggggtgggga
54061	gtggggcact	gtggccccag	gcaggttctc	tccccacagt	gctggggctc	ggggcctggc
54121	cttgtgtcctg	gaatcacaca	tgacagggtg	ggggagacag	gggagctaat	gccatttgcc
54181	tgctctgcat	ctctgttctc	gagaatggcc	aggtccccct	tcagcagctg	cttgtgtggc
54241	ctgtgggggaa	ggaaaggagg	tgaggttgtc	ctcatctcca	cggtcttggg	ccccctctcc
54301	ctccccactc	ctcgaaggaa	cagggtctgt	cttggccgcc	atgacagatg	agaaatactga
54361	ggctcaaaag	gggtgagcag	cctgctccaa	gtgacacgat	gacaaaagac	gtgagttctga
54421	atcaaatgtg	tctgctgtt	gctccacctc	acccaaggca	gctggaagtg	gttagaaacg
54481	cacgttctca	ctggagagag	aagggtcctg	gagaggcgag	gtttggcagg	agggccccgg
54541	gccacatact	tatgttggcc	aggcagcttc	caggctcagc	ctcgggctct	gggtctctcg
54601	cgaaatgagc	ctgccagtc	aaactgctga	ccagctctc	ataggccagg	agggccgtga
54661	agaccgcggg	ccgtggcggg	cccttggcaag	catctccgaa	gctgtgcagc	ccggccacag
54721	accatgtcgc	ccctacacctc	tgacaccagt	gtgccccaga	caggctctcag	ctagggtcag
54781	gtgacactgtg	gtgacttttt	ataggcagct	gtgacaggag	gttagagcag	ctaggggaaac
54841	ctctgctgtt	tggggggctaa	ttggcaaaaa	ggcttagttt	caggttggag	gtgggtccag
54901	aggccaagcc	tggaaggttc	gttcttgagc	tattgggtga	ggggatgccca	ggggccttgt
54961	aggttaaggt	tgggggccgt	gggctcacct	cacagctggg	cagctcaccc	acagcactgt
55021	tacacacact	ccccggcaga	atagggtctg	catcaccccc	aggagctcga	tgacggcgcg
55081	tgcaaggccct	aggccccag	agggtcacgg	gcactgtctg	gacagcagcc	atgctgtgtg
55141	agcaagggaa	agctgtgctg	cccgccctgc	aggttggatg	gacagcagcc	ctgcgccctg
55201	cgccacactc	ctgctcctgg	gcggggccgt	cccgagaccc	agccacagct	cccatcaagc
55261	aggtgtgtgt	cagSataggg	caggcagagg	ggccgcaggg	tggtctccag	gtgtcaacgg
55321	gtggccagca	caggaggggc	catgtcgtat	ccccctccag	ggtgggtgta	ggtctcagtc
55381	aggatgagct	gcttcaggcc	ccactcctcc	gggtctgtcc	ccagccctac	gctccatctc
55441	cttggggcct	ggcgccctgt	cagaggcagt	gtggacggca	ccaggtgaca	gggtctccgg

55501	caggggcagg	gggcacagca	gagagggatc	caagactcac	ccaatgaagc	agtgggcagc
55561	agttagcacc	gctctctctg	acaccagggc	tcgcccaacg	ccagctgcat	ctgtgtgcat
55621	cagctcgccc	tcaccaggcc	atgggagagg	tgctctcgcc	tggggagacctg	ctgtctctcaa
55681	ggatccacag	gctgaaacag	ataagtgcct	catctgacat	cttgttaagc	acttcccact
55741	tcocatccag	tcctcacagt	agcttttgaa	attgaatttt	gttcccctca	ttaactgaag
55801	taaccaaaag	tcagagagag	ttaaRagact	tccttgctct	gtactccctg	cgcaagtggca
55861	gaacttgcag	tcgcaagcca	gggcctgaag	ttttagccct	gttggtctga	tgggacccct
55921	gtcccagctc	ctgtgggaatg	cccctacttg	tcactcccct	gtctcaccgc	ctgaccacct
55981	ccaggtMtg	gctgcctccc	gtattcctga	ctcatttagt	catctcttat	catcaagaca
56041	taattctcgt	cccaggaaat	actttgcatt	caatttgcat	gtactccact	taatgcgca
56101	attcttgactg	ggagccctac	cctgtgcagg	ctcgtcgggt	ccctgctgga	agcctgccac
56161	ttcccacagaa	acccaagtca	ggtctcagag	attcctcttc	tcaccttaac	tcaccaacctg
56221	tagagttcca	aagtgcctgt	gcctctcagc	cctaacaggg	ctgttcccat	cccagggggt
56281	gaagagacct	cctaattag	ctcggcggtg	tggatgccta	tgccagctct	ctagtctcaa
56341	ctgaggttgc	cttcacagtg	gcttctgccc	actcccagca	tgcccacatc	gtactctaac
56401	agctgtctctc	atactcatc	tcgggggtct	ctgggctctg	ggcgagaaa	gctgcccct
56461	gaactcgagc	ctgcagccag	gaactgtgag	cagctgtgtt	gggtcagcag	acagggagct
56521	ctcctctggc	acagcttgat	gcnaagctga	tgatgcagc	gtgaaccag	tgctcctcag
56581	ctgcagagca	cagcacagg	ccccgggaat	ctcctggag	ccagccaaca	acagcaagga
56641	cagatgcctg	agccagctta	tgccagacac	ctctgtattg	cagStcttct	ccccagctg
56701	gtgagtccaa	ccccagcccc	agggttaggc	ccctcccctt	ctgctctctc	tcttctccct
56761	atcacagactc	acaggggccc	tgaccccag	gctgggggccc	cccaacttc	atcccaggcc
56821	gggcccgggt	ggacaggtgt	cgctgtgtca	gctggttgt	gatacagta	catgtggggc
56881	catgtgatag	agccagcgcc	agattgcgta	gggtcccagg	agctggtgaa	agcagcgagg
56941	ctggggctag	agtcctggat	ctgggaagca	agggaagact	gaagccagga	cagttvggag
57001	ctgagactgt	cagcctgggc	aagcagagtg	ctctccagc	cttcatgag	ggccagcaat
57061	taccatcact	gggtctctga	tcaccagcag	tgcccacga	ggaggtctca	aagggggaagc
57121	gatggcgagg	ctggggcagg	cagaggggtg	tggtggctgt	ggcggtgggc	agctgcagca
57181	gggcccagctc	tgagccctgg	ctgtagtgtg	tatagccct	gggcaactgc	agggcagcca
57241	ccccacagtc	tctggcccac	gggtgagtc	ctctcagctg	cagagaaccc	agatagcagta
57301	accaggaatt	cagttctgtt	gctgctgccc	tggaagggga	aggacaagga	ccaggtctgt
57361	gagtgcaagg	gtagacagtg	acaccatggg	agaccccaca	ggttagtgta	cagatgacgt
57421	gagcactcca	cggttttttc	tttttgaggt	agggtctctg	ctgtctgccc	aggtctgcag
57481	cgactggcat	gacctgact	cactacaagc	cttgacctcc	cagctctcaag	cgactctccc
57541	acctcagctc	cgtgagtgc	tgggactaca	ggcgtgcacc	acctgctctg	gctaatttta
57601	cagtttttgt	agagacatgg	tcttgctatg	ttgcacagcg	tagtcttgaa	ctYttggSt
57661	caagcgatcc	tcctccctgt	gcctccccaa	gKgtcgagat	tacaagcgtg	agctacttgt
57721	cccggctcca	gtctgtgact	ttaaagcgag	tcattccctt	gcaacctgag	ttttYtgatt
57781	tgtaaatctt	ctaccatgaa	agggtgttgt	tacaattaac	ttgacaatat	gtaccagta
57841	catggcagat	tctcagtaaa	tggttcttta	ttattactct	agaggcagg	actgaccttg
57901	tggggcttga	gacaaattgt	ccactcccca	gagccggaga	ccctgctgta	ccccagctg
57961	gacttacctt	tcaaaagcag	gggcagcag	gaggaacctg	gtgtctgcca	cacagggaac
58021	gctgcagatg	tggtctcctt	gcctctctac	actggctcgc	cagggccaat	cgccagggaac
58081	tggtgtgccc	tctgaggtc	tgggggggcc	ggggccacgc	gtgccacag	ctgtggaaag
58141	gagttagtca	cactgacagc	agatacctgt	cctgacctca	ccccagctcc	aaaccagacc
58201	cagataagac	ctaggccccg	tgtccccaga	ctttaccacg	ctgagcgctg	tgaggagggag
58261	ggagagagga	gacacaggt	agatgcagg	actccagccc	tgctctagctt	tgggggagag
58321	aggggatggg	ctggggggcg	ggcccagggt	ccctgctgtg	agctctgatt	accttaccag
58381	ccccctccag	aatgaaaata	tttatatag	gcgagacag	cttttatgg	aaacctatca
58441	gtgtgcacac	ctagtaatta	attctccttc	agctgtgctc	agcaactcgg	tggtgggggt
58501	tcagcttggg	agtgaggcca	ccctctccca	ggcttcagaa	cccccaactc	gtgcccctgc
58561	cactgagctc	gcagggcgcc	ctgtgtgtgt	agagagcatt	agcttaattg	ctctcttagc
58621	agcagaggcc	taagaggga	gatttagagc	ctgcatcatt	tccaagtggt	gagggcccca
58681	agaaaatgag	acttacctca	ccctgtgcta	gagactcagt	cttcccacac	tccccagaaa
58741	ctgtctgaga	gccccccaga	gagaggggcc	ctgcgccacg	ccctccacag	gcacacagc
58801	accocatgag	acagctgagc	caggctgccc	agagagtgga	tgaaaagaaa	gggaactgg
58861	ggccagagga	gcccagagtt	tgcttgagct	cattgtggaa	gtcgaggggg	agggcaggac
58921	aggaacacag	caggcagcac	ctcacacaca	cacagccccc	gagcccccgc	caacgcaaac
58981	tgagctcggtg	ccgagaaaaa	ctcatccatg	ttgacacag	tgggcacata	taccacccca
59041	cagacttcta	cgaggttaca	tcccactggc	tggaactgtc	ccccacagc	gtccccccca
59101	caaggaagcag	ggactgctgg	ggcagagatg	gcccctgagc	ccccacagc	ccacacccat
59161	accccagcag	atggcggtct	accctccatg	aggactgtgg	acccccgctg	gagcagcact
59221	gggccccagc	accacttcat	gctgccccgc	gcactctgct	ccactgtgct	ccactgtgag
59281	agagggccac	tggttctccc	tggtctccac	ctcgtccac	ctccagttgg	ctagggattca

59341	gctgtctgtct	tgccctagcc	agcagtttct	agctctgggg	ctgtgtgtgt	gtgaccctag
59401	gctgtctctc	acccctctct	gggccttagc	ctgtgtgtgt	gctgtgtgt	gctgtgtgt
59461	gtttccttac	tacttgagtg	gctgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
59521	agtggtgaa	tggaagagc	ctcaatgac	gctgtgtgt	gctgtgtgt	gctgtgtgt
59581	tgagagtg	atcttggaac	catgtggagc	agctgtgtgt	ctctctccac	ctccctctct
59641	tcatttccac	tcctccactt	ctctctccac	aggggttagg	tggtgtgtgt	atgctctct
59701	ccaggtctgc	tcagagagac	tggtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
59761	gtgtgtgtgc	ctgtgtgtgt	tggtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
59821	gggaagagc	gtgagagagc	tggtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
59881	gggaagagc	gtgagagagc	tggtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
59941	agagcattca	gggaagagc	tggtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
60001	tatagatag	tggtgtgtgt	tggtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
60061	aagcagagc	atcactgtgt	gtgtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
60121	ctgtctctac	taaaaatca	aaattgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
60181	tactgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
60241	cgaaattgtg	ccattgtgt	ccagcgtgtg	caacaagagc	gaaagtcgtg	gctgtgtgt
60301	aaccaaccaa	aaaaaagca	tatagatgt	ggagccagat	ggagccagat	gctgtgtgt
60361	ctgagagcct	ctgtgtgtgt	ctgtgtgtgt	ctgtgtgtgt	ctgtgtgtgt	ctgtgtgtgt
60421	atcatcaaga	taggtgtgt	aaatgtgtgt	aaatgtgtgt	aaatgtgtgt	aaatgtgtgt
60481	gttttaacag	tttaacacag	ttttttttgt	ttttttttgt	ttttttttgt	ttttttttgt
60541	ctgtgtgtgt	aggtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
60601	caagcagct	tcctgtgtgt	ttccgtgtgt	ttccgtgtgt	ttccgtgtgt	ttccgtgtgt
60661	ctgtgtgtgt	ttttttttgt	ttttttttgt	ttttttttgt	ttttttttgt	ttttttttgt
60721	gctgtgtgt	aggtgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt	gctgtgtgt
60781	gattacagct	gtgagcagct	gcaacgtgt	gcaacgtgt	gcaacgtgt	gcaacgtgt
60841	taggtgtgt	ctgtgtgtgt	aaatgtgtgt	aaatgtgtgt	aaatgtgtgt	aaatgtgtgt
60901	tcgaagagc	ccaatgtgt	ctgtgtgtgt	ctgtgtgtgt	ctgtgtgtgt	ctgtgtgtgt
60961	aactcaagga	agaaactgt	agcaagcatt	agcaagcatt	agcaagcatt	agcaagcatt
61021	aggtgtgtgt	tcagcctgt	caagagagga	ggcctgtgt	ggcctgtgt	ggcctgtgt
61081	ctgtgtgtgt	gtttctgtgt	gggtgtgtgt	gggtgtgtgt	gggtgtgtgt	gggtgtgtgt
61141	caactgtgt	gagatgtgt	ataacgtgt	ataacgtgt	ataacgtgt	ataacgtgt
61201	ctcaacgtgt	ccagctgtgt	agaaagcag	agaaagcag	agaaagcag	agaaagcag
61261	cgagacacat	ggttcaagct	tggtgtgtgt	tggtgtgtgt	tggtgtgtgt	tggtgtgtgt
61321	ctgtgtgtgt	caatgtgtgt	agctgtgtgt	agctgtgtgt	agctgtgtgt	agctgtgtgt
61381	tggtgtgtgt	aaagagcag	ctgtgtgtgt	ctgtgtgtgt	ctgtgtgtgt	ctgtgtgtgt
61441	aggaagaggt	ggtatgtgt	agcgtgtgt	agcgtgtgt	agcgtgtgt	agcgtgtgt
61501	cccttaggca	aggtgtgtgt	gcaaaagca	agcaggtgt	agcaggtgt	agcaggtgt
61561	gggtgtgtgt	ctctgtgtgt	tgccgtgtgt	tgccgtgtgt	tgccgtgtgt	tgccgtgtgt
61621	tcaggtgtgt	gtgtgtgtgt	taggtgtgtgt	taggtgtgtgt	taggtgtgtgt	taggtgtgtgt
61681	agaaacagat	ccaggtgtgt	tagagagag	tagagagag	tagagagag	tagagagag
61741	tcaggtgtgt	ggccagcgt	gtccgtgtgt	gtccgtgtgt	gtccgtgtgt	gtccgtgtgt
61801	ggcttcaagc	agcgtgtgt	tcctgtgtgt	tcctgtgtgt	tcctgtgtgt	tcctgtgtgt
61861	agaggtgtgt	gggaaggtgt	tcataaagca	gaggtgtgt	gaggtgtgt	gaggtgtgt
61921	cgactgtgt	aaacactgt	ggaggtgtgt	ggaggtgtgt	ggaggtgtgt	ggaggtgtgt
61981	agcaggtgtgt	gggcaacata	gaaagagcgt	atctgtgtgt	atctgtgtgt	atctgtgtgt
62041	ggcaggtgtgt	ctcatgtgtgt	taatccaggt	atctgtgtgt	atctgtgtgt	atctgtgtgt
62101	tgaggtgtgt	agttcgagag	cagcgtgtgt	aaacgtgtgt	aaacgtgtgt	aaacgtgtgt
62161	tacaaaaagt	agcgggtgtgt	gtgtgtgtgt	agcgtgtgt	agcgtgtgt	agcgtgtgt
62221	gggtgtgtgt	cgctgtgtgt	taggtgtgtgt	aggtgtgtgt	aggtgtgtgt	aggtgtgtgt
62281	catccagct	tggtgtgtgt	gagcaagct	tcgtgtgtgt	tcgtgtgtgt	tcgtgtgtgt
62341	tgactcaagc	ctgtgtgtgt	agcactgtgt	gaggtgtgtgt	gaggtgtgtgt	gaggtgtgtgt
62401	gagattgtgt	ccatctgtgt	caacatgtgt	aaacgtgtgt	aaacgtgtgt	aaacgtgtgt
62461	tagctgtgtgt	tggtgtgtgt	cgctgtgtgt	ccaggtgtgt	ccaggtgtgt	ccaggtgtgt
62521	atgtgtgtgt	ccgggtgtgt	agaggtgtgt	gtgaggtgtgt	gtgaggtgtgt	gtgaggtgtgt
62581	ctgtgtgtgt	gagcgtgtgt	atgtgtgtgt	aaacgtgtgt	aaacgtgtgt	aaacgtgtgt
62641	tggtgtgtgt	gacatgtgtgt	tagtctgtgt	tactgtgtgt	tactgtgtgt	tactgtgtgt
62701	agtgagctat	gattgtgtgt	ctgtgtgtgt	gctgtgtgtgt	gctgtgtgtgt	gctgtgtgtgt
62761	aaaaaaagca	aaaaaagc	ttcaacaggt	ccctgtgtgt	ccctgtgtgt	ccctgtgtgt
62821	tgccattgtgt	caggtgtgtgt	aagtgtgtgt	caataaagt	caataaagt	caataaagt
62881	ctgtgtgtgt	gtgtgtgtgt	ggacgtgtgt	gctgtgtgtgt	gctgtgtgtgt	gctgtgtgtgt
62941	gagctgtgtgt	gggtgtgtgt	aaggtgtgtgt	ggaggtgtgt	ggaggtgtgt	ggaggtgtgt
63001	agctgtgtgt	ccgtgtgtgt	gcactgtgtgt	ggcctgtgtgt	ggcctgtgtgt	ggcctgtgtgt
63061	tggtgtgtgt	ctctgtgtgt	atgtgtgtgt	atataaagt	atataaagt	atataaagt
63121	gacaaaagat	agctgtgtgt	cagctgtgtgt	ttaaaaagga	ttaaaaagga	ttaaaaagga

63181	aaactggggg	gttggtgggg	catggtagct	cacgcctgta	atccccccag	cactttggga
63241	ggctgagggc	gacagattgc	ttagagaccg	cctgggcaac	atggcgcaac	accatctcta
63301	ccaaaaaaa	acaaaaacaa	aaatlagctg	ggcataagtg	tgcacgcctg	tgtattccag
63361	tgcttgggag	gctaaggttg	gaggatccct	tggcgaggga	ggcgagaggt	gccatgaact
63421	gagatcacgc	caagtgcacac	taagggcctc	ctagacctca	ctttgggcga	taagaccaga
63481	ccctgtctca	aaacacaaac	aaacaaaaaa	cctgggggac	taggatgtct	tttagggccc
63541	ttcagctctc	aacagctactt	aaaccaatta	aaagactctc	gttagttacc	ccccacatc
63601	cccccccgca	ggacgctcSg	tgatgagcag	ctagctggct	gtcacgtctg	tggatcacca
63661	agattgcatg	ggatggggct	gagctgacca	agggggatga	ggggccgggg	ggggcgggca
63721	gggagggggg	ggagccactc	acctaaacat	agctgaagt	tgtagaagtg	gcacccgaat
63781	atgtgttgg	attgattgag	gatgtctgtc	tgtccacgca	catgtctcac	cagccccaaa
63841	ccctgcccc	acctggcaga	gggggggggt	gggggtgaac	caggttagga	ctgtcaacc
63901	agtgctcttg	acctggcccc	agaaaggtga	tttccaagaa	gccactctgg	ctatctcttg
63961	tccccgacc	tccccactca	gtccaaaggt	cgatgatctc	ctggaccagg	gcacctttgg
64021	ccacgtcaag	attccatgtc	actgacctca	tctctccctc	tccccagacc	agggccggac
64081	gtggctactc	cgtagggcct	gcttttctc	ttagaacctta	agtaaagctc	tttttttttt
64141	tttttttttt	tttttttttg	acggagttct	actctggccc	agggctggat	cgagtggcgc
64201	gatctcggt	cactgaaacc	tccgcctccc	gggttcaagc	gtttctctgt	ctctagcccc
64261	cccgagttag	tgggattaa	gcaccccgca	tagcccccga	ttaatatttt	tatttttggt
64321	agagacagtc	tttaccatgt	tggccaggc	tgttcaactc	ctgaactcaa	tgtatccatc
64381	cgctctggcc	tcccaaaagt	ctgggattac	agaggtgggc	caccgcgcgc	ggcccttaag
64441	taattcttaa	aattggcaag	ctggataaac	ggttcactcg	gttttgatca	agaagactgg
64501	agtcgggggc	agattattctt	tgccctggac	cccgaaatct	ccagctcccc	ggccactcac
64561	tcgctctctc	tgtattccgt	cattatgcta	acgctggccc	atcacgcaac	cgcacactgc
64621	gccaccttgt	tctctgggcg	agccatcgcc	aaaccccccc	tttccctcgt	cgccctcttt
64681	gagaccatct	tcaatctcta	cgccactcct	gctctccgcg	ctttctcgtt	acgcgttttt
64741	ctcttggtact	caattcacgt	gcgagtcctc	ggaataatcc	cagtcoccca	cactgtctgtg
64801	tccttctgct	cgactcttta	tttcgaacac	caagtatccc	gggttagcag	gcocctgtgc
64861	aacgaccocg	cgagcagttc	agccccgtgt	ccgttccccg	ggcacaacga	tccccagact
64921	cagataaatt	atctggcatc	ctggccgcgc	tgtctccgag	ccccccagtc	ccccactccg
64981	tgcacacctg	gaggagaaga	cgccggaaca	gctgatggcg	gtccccagct	cgacagcgcg
65041	gcggtaattc	cgctccccgg	cgccgcgcgc	cttcacgtgc	cgccgctaga	cgagagagag
65101	taagcccgct	agggcaagag	cgagcccgac	ccagccaagg	ctccccaggt	tgtcgccatt
65161	tatctccag	tcccgcgcga	ggccgcgcgc	gagaaaaacca	gcccagagac	aggggcccgg
65221	cgccgaatgg	cgccgccccct	cctggccctc	tgaactcgcg	attggccggc	cgctctcgca
65281	ctccaccgac	caaatggcgt	ttccagggcg	ctagtcaagc	ggccagattt	ggaaaaacag
65341	gaagaatgcc	gggactagtg	aagccggtaa	gggacgtgcg	gaatcgcgcc	ccccagcgct
65401	gccaggcatg	atgggagttg	tagtcggcgc	ggctcgaagg	catcaaggga	aatgaagttt
65461	ccacagattt	aaaaactgtt	ggccgggca	ggttgctcac	gcttgaatc	ccagcacttt
65521	gggagccgca	ggctggcgga	ttacctgagg	tccagagttc	aaagaccagc	tcgccaaact
65581	ggtagaaccc	catctcaact	aaaaatacaa	aaatttagcc	ggcgtgtggg	cacatgcctg
65641	taactccagc	tactcgggag	gctgagggag	gagaattgct	tgaagccggg	agccgagagt
65701	tgcagtgagc	cgagatctgt	ccactgcact	ccagccagcg	cgacagagtg	agactctgtc
65761	tcaaaaaaaa	aaaaagaaaa	aaaaagttgt	ttagttgggt	taacagagtt	tgcgcttacc
65821	ctatgcccaa	tctgtttgta	agaattgcag	catccgggac	ctagagacca	cgcgatcagg
65881	ggatccagca	aattacggcg	tcgagttcgg	gaaccaagca	tttcccccta	aactatttca
65941	ggcaccaatt	gggctgcagc	ctccactcct	cccggttctc	gcctcaaccg	tgtcttccat
66001	tggctggctc	ccctttctcc	catacattca	cagaaccact	ccittggcca	cacacacctc
66061	tgcagtcatt	ctaacctctc	tactatcatg	gcgcccgcct	ggcaaccact	gagcttgcatt
66121	caactcaaca	ctcagttctc	cctttccaag	ctgtggcgca	cttagaggtc	ctccttcaca
66181	gacgtcccat	ctgatgttga	cccatctctc	ctacaccttc	aaactctcca	tcttccccat
66241	ctactccagc	aottgtttct	ctttccacct	tcagaaaacag	caacaaattc	ctcttcaact
66301	agaggtctct	caccagcctc	ctctcccgcc	attatcccat	ctaccctctc	acatttcaagt
66361	tttttgaaag	attctacact	ccagctctct	acttctctac	tcttctctcg	ctgccacagc
66421	cataaactcg	ctgtgcctct	cagcatttgc	ctgcacacta	gtggctgtgt	ctccaaagac
66481	cgtagaccga	actgttattt	atttatttatt	ttactatttt	ttagacggag	tgcactatgt
66541	tcgcccagcg	tggagtgtag	cggtgccatc	ctggctcact	gcaacttcgc	ctccagaggt
66601	tcaagtgtgt	ctcgtgcctc	agcctcccaa	gtagtttggg	ctacagctgt	ctgccacagt
66661	gtctggctaa	tttttgtatt	tttagtagag	acagggtttc	accatgtgtg	cagaggttgt
66721	cttaaacctc	tgaacctca	tgtatcaccc	acctcgccct	cccaaaatgt	taggatttga
66781	ggcgtgagcc	accgcacccg	gccaatgtgt	gtttttcagg	tcttctctct	cttgactctc
66841	cagagggatc	ccttactgtt	gcacctaccc	tctctgggaac	tcttctctct	tgcgctgtct
66901	gatatttccc	tctctctgtg	gctctccctc	ctccagatgc	tgtttctcac	atctactctc
66961	ttctagagag	tgtgtagagc	agaataatgg	tcaccaaga	tgtccctgca	tgaattccct

67021 gaactttgga atatgatagg ttaaatggcc aaaagggaa taaggttgca gatggaatta  
67081 aagctagccaa tctcctgatt ttatttttatt ttattttggt tttaggttgg agtttctctc  
67141 ttgtttgccca actgagtagtc aatggtgtga tctcggctca ctgcaacctc ccgctgccag  
67201 gctcagagaga ttctcctgcc tcagcctccc gtagtctggg gattacaagg acccgctgat  
67261 atgctctgact aatttttttaa atttttagta gagacagggg ttctctalat tggccagggc  
67321 ggtcttgaac tctgacctc aggtgatccg cccacctcgg cctccaaaag ctgctgggatt  
67381 acaggcggtga gccaccgtgc ccagcctatc tatctattta ttattttatt tttagatagg  
67441 agttttgtct ttgttgccca ggctagaatg caatggtgct atctcgactc acccgcaact  
67501 ccacctctccc ggttaaaggcg attctcctgc cttaggctccc tgagttagctg gggattatagg  
67561 catctgccac cagcctgggc taattttttg ttatttttag tttagatggg ttttccctagt  
67621 ttggtcaggc tgggtctcaa ctctgacctc cagatgatcc acccactcgg gctctccaaa  
67681 gtgctgggal tataggcggt agccatcata ccaggctcta ttgatttatt ttatttttta  
67741 tttttgagac gtagtctcgc tctgttgctc aggtcggagt cagatggcac aatcttggtc  
67801 cattataact tccgcccccc cccaggttca agcattctc ctgcgccagc ctcccgacct  
67861 gttgggacta caggcgctgc caaccatgcc tggctaaatt ttgtattttt agtagagact  
67921 ggggttccaat gtgttgccca ggctgggtct gaactcctga ctttggtatc tggcctggc  
67981 agcctcccaa agctgtggga ttacaagtg aaagccaccac gccccagcta tttgtttat  
68041 tttttccaaa acccttgaca cccaggtcgg agtgtagtgg cactgtcata acctcaatgca  
68101 accctcgtct cccaggttca agcatttctt gcaactcagc ctcccttagta gctagagtaga  
68161 caagtacgtc ccaccacacc ttgctaaatt atttttatt ttgtagagat ctggctctcac  
68221 ttgtttcccc aggtcgtgctt aaacttctgg ttccaagcaa ccttcccacc tcaaatggtt  
68281 gggagtagag gcatgagcca ccaccacacc tggcctaatt tggctatttt taatttttct  
68341 ttattatttta tcttaatttt taatttgaga cagagtcttg ctctgtcatc caggctggag  
68401 tacggtgttg caatctcagc tcactgcaac ctccccctgt cgggttccaag cMatctttgt  
68461 gctcacaact cccaagtagc tgggattata ggttctggcc accacgcctg actaattatt  
68521 ttaattttttt ttgtttttag tagagacggg atttaccat gtgtgccagc ctggtcttga  
68581 actcctgacc tcaagtgaac cactcgtccc agccttcaa agtatgggga ttacgggtgt  
68641 gagccgcgtc cctcgcccca atttttgat tttagtgga gtaggggttt taatgaagga  
68701 gccagcgctg tctggaactc ctgacctcag gtgacccgct ggtcctccgc tctcaaatgt  
68761 ctggcgttac aggcataagc caccatgcct gcccacag ggtccttaaa aactgaagga  
68821 ggtgtgcaga agaaagtcaag agggagatgt gtagtaagaa aaaagacaca gagagctgca  
68881 atgtttctgg tttagaatg taggaagggg atttgagct aataataatc ttgggctct  
68941 aaaggccaaga aagggtaaaag aactggattc tcaacttaga gtcaaccggga aggaactatc  
69001 aacatcttga ttacagccca gtgagactct gtgagactct taagctacag aactgtatgc  
69061 taaattttgt ttgttttaca tcatataatg tgcagtaact tgttacagca gcaattagaa  
69121 atgaatcacag aggactgggc attagccctg tatctcagct tctctcagc tctgtgtgtg  
69181 ttctctgtat ttattgttgg ttccccag aatgagtgat ctaaggagaa gcaaaataga  
69241 agccgcaact tctttatgac tttagcctcag aagacacaca ctggggccag tgcagttgc  
69301 ttatgctatc aatccagca ctttgggagg ctgagggcagg aggaccactt gagccaggga  
69361 gttttgagaca cctgggacaa cacagggaga cctcactctc aaaaaataa aacaaaata  
69421 gccaggtgtg gtggtgcaca cctgtagtcc cagaactttg ggagctctag acaagacaa  
69481 gacttgagcc caggagtttg agacaggtct ggacaacgtg gtaagactg ctctttata  
69541 acatttttaa aataggcgg ggcatggttg ctcatgctg taatccagc aatgtgggag  
69601 gctgaggttg gtggtacacc tgaggtcagg agtttlaagc aagcttggcc aacatggcga  
69661 aaccYcgtct ctaactaaaa Yacaaaaat agccgggcat ggtgtgtggt gactgtatc  
69721 ccagctactc aggaggtcga ggacaggaga ttatttgaac cgggaggtg gaaagtgtag  
69781 tgagccgaga ttgccttctc cactccaaga gttataaagg attttgacca taattttgtg  
69841 tagcaattaa tcaatttaatt aattlaatgt agacagtcct actctgctgc cactgctgaa  
69901 atgcagtggt gcaatctcgg ctcaactgaa cctctgcctc cggagtcaa gtgattcttc  
69961 gtccttagcc tctgagttag ctgtgattac aggcaccagc cactatcgtt ggtctattt  
70021 tgtgttttta gttagagcgg ggtttcacca tattggccag gctgttcca aactcctgac  
70081 ctcatgacc Rccctccttg gcttctcaaa gtgctgggat tacaggctgt agccactgtc  
70141 cctggccttt ttcttctctt ttttttttt ttcattagag atagagtgtt gttgtgtg  
70201 ctctaacttc tgggctcaa gacttctccc accctgctt cccaagtgc gctgggatt  
70261 acaggagtag gccactgccc ccagcctctg acagtttttg tgcactagga attttgggag  
70321 caaattttac ctggctattt ctggctcata caattgcagt cagatgtggg cttaagctgt  
70381 acaataaagc agctaaaaa gctgaaagat aacctagcat tctctctccc tttctctgag  
70441 taagtctcga acctatctat tgtgtcctct gctatgggcta gcttgggct cctcacaga  
70501 tggcagcctt aaggctttaa tagtcagctt ccaaaatggt cctcagtagt tctctctc  
70561 tggatttag accattgtga agtctctctc tagaggtcaa atcatgaaag tttgtggct  
70621 ggttaatatga gaaatgacag tgtgtgactt ctgacaaaag ccaagtcaca tatgtatgct  
70681 tccaactctgc tctcttgtag atcaactcatt cttagtagag gaactgaggg cctgtaagaa  
70741 gcaactcaag aaccttaggg agagaggtct ccttagtagag aactcacaag cgtgactagt  
70801 acgtgggtgt gctgcagtca agtgggcata

70861	gagtttagat	tgacggcata	tagctccact	tgttatcaca	ccgctgtagc	cataacatgg
70921	gaaggctcat	cacttggctc	tgagccactg	ttgtctgtaa	agggataaat	gccctctgtg
70981	acactgtgcca	cagggtcggg	cccaacatgg	cttgacatgg	gacataggtc	ttgtgcaggt
71041	gcttgtacac	agagaaagag	agaaagccag	agctgtccat	ctcgggagat	caacagacaca
71101	gctcataatg	ctcatggcca	gagggagaaa	gagtaaggct	gtgggggtgtg	gtggctcatg
71161	cccgacagcc	gcgacactttg	ggagggccag	gcaggtggat	cacagaagttg	gaggtttgag
71221	accagcgctg	tcaacatggg	gnaacccctt	ctcaactaaa	aatacaaaaa	ttagctggcc
71281	atgggtgtgc	atgcctgtaa	tccagctatc	tccagaggtg	gagcgagagc	aattgtctga
71341	accagagagg	cgagggttgc	agtgcgcca	gatcacacca	ctgcactcca	gcttgggcaa
71401	cagcgcgaga	ctccatctca	ggaaaaaagg	aaaaaagaaa	agaaaagagc	gaagtcgtga
71461	gctgaccctg	aaggggagag	tgcccaacaca	gctgtgtgtg	tgtggagact	gccggagtaa
71521	gcagctgaga	cgagcagcac	agtcgagag	taagatgtg	atgatgagag	agctgctgaa
71581	taaaagccat	tctcatttac	ctgctgtctc	tcgagtgttc	tctcagctcc	ctgctccacg
71641	tcacactgtg	ctctcaacac	ctcagctggg	gctggaccoc	aacctgtgac	atgcggggcc
71701	ttctgtcaac	aaccagcagt	aacctgctgg	gcatgtgagg	gagctacctt	ggaatcagat
71761	tctgtaaaa	agtcacgctt	tcagatgacg	gtacatttgg	ccaacatttt	gactgcactt
71821	catgagagac	cctgagccag	aaccccttag	attcctaacc	caagaaaact	tgttztgata
71881	agtgtttatt	gttttttttt	tttttttttt	tttttgagaa	agagtctcgc	ctctattgcc
71941	agggctggagc	acagtgccac	aatcttggct	cactgcaagc	tcgcctctcc	aggttccaac
72001	cattctctct	aatcagctct	ctgagttagt	gggactacag	gcacccacca	ccacgctcag
72061	ttaatttttt	tgtattttca	gtagagacag	ggtttcaccc	tgttagccag	gaggtctcca
72121	attctctgac	ctcgtgatcc	gcccgctctc	gctctccaaa	atgctgggat	tacaggcgctg
72181	agtcaccaca	cccgccacgt	gtttatttgt	taagatat	ggctaggcgc	agtggttccac
72241	taacatgata	gcgcactttt	gggaaggcca	agtgggagga	ctctaatcgc	tcaggagttc
72301	aaagtccaaga	acagcctggg	caacatagtg	agaccttgtc	tctattttaaa	aaagtgtttt
72361	taagatgtta	tcttctagct	gggtatgggt	tggtctacgc	ctgtaatcgc	agcactttgg
72421	gaggtcgtag	tggtgggact	acctgtggct	agagatgga	gaccagcgtg	gccaacatag
72481	tgaacacccc	tctctactaa	aaatacaaaa	aattagctgg	ctgtatgttg	gaggttaatt
72541	aatccccagt	actagggagg	ctgaggcagg	agaatcgctt	gaacccggga	ggcagaggtt
72601	gcagtgagcc	agaatcgtgc	cattgcactc	cagcatgggt	ctatgttttg	gaggttaatt
72661	gttaccacag	aataaataat	tctgacaggg	caccagcctg	gccaacatgtg	agaaaacccc
72721	ctctagtaaa	aattatccgg	gtatgggtgt	gcatcgctgt	aatccacgct	acttggggag
72781	ctgaagcagg	agaatccctt	gaacctggga	gggtggaggt	cgagttagcc	agaatctcac
72841	cactgcactc	cagctgggtt	cacagagcaa	gactctgtct	caattttaaa	ataaataaat
72901	aataatatag	ggcagtcaga	ctgccacact	ggcagctcag	gactagcaca	tgtgtctcag
72961	aaagccaggt	ggaagctaca	tattttatga	tctaaactca	gaagtcagat	agcaatctgt
73021	ccactgtaat	cacaagcctt	cccagttcca	aggggagggg	acatagcctc	cctcacctct
73081	tgatacaaga	agtgtaaaa	ttatatggta	agaagttggc	caggccctgc	ttgtctctgt
73141	tgttcatgcc	tgtaatccca	gcactttggg	aagacgaggg	agatggatca	ctcaggttca
73201	ggagtttgaa	gccagcctgg	ccaacatggg	gaacccctat	ctctacaaaa	atacaaaaaa
73261	tagctgggca	tggtgggtat	cacctgtaat	cccagctact	tgggagggca	agaccagaga
73321	atttgttgaa	ctggggaggg	agaagttgca	gtgagccgag	attgtgccac	taccctctgc
73381	cctgggttat	agtgcaagac	ctctgtctaa	aaaaaataaa	aaaaaataaa	agagaagttg
73441	gtggggccca	gtggctcacg	cctgtaatcc	caacactttg	ggaagctgag	atggggaggt
73501	cgcttcaggc	cagaagatcc	atcgtttacca	gcctgagcaa	cacaagtgga	tcgcctctct
73561	acaaaatttt	tttaaaaatc	agctgggtgt	gggtgcaggg	acctgtgata	acagatcact
73621	gggatgctga	ggttaggagga	tcgcttgagt	cagggaggtt	gtggctgcgc	taagccatga
73681	acatgcctga	gcattctagc	ctgggttaaca	gagtgagaca	ctgtttcaga	agcaataata
73741	aaataaaaaa	aataatgtgt	taggacaggg	gtggggctca	cgcttggaat	tctagttgct
73801	tgggagactg	agggcaggag	atttgttgag	aacagcaggt	cgagcgtgca	gtgagctgct
73861	atcgccaccac	tgacactccag	ccttggtgac	atgagcgata	tcttgttcca	ataaataaat
73921	acatacagtt	ctcttttaca	tgaagtata	gtaaaatttt	aaaaatacat	atggggaggt
73981	tagaaaagccg	cctgactctc	cctctccctc	tcctctccct	tctctccctc	cctctccctc
74041	tcctgtctcc	tctccgtctc	cgctccctgc	tcctctccct	gtctctctcc	cgctctctcc
74101	tctgtatgcc	agccaaggct	ggacggtgct	gctgccatct	cggtctcagc	cagctctctc
74161	gcctgatctc	cctgcctcag	cctgctgagt	gctgcgattt	gcagcgcacg	gccgcacacg
74221	ctcactgggt	tctgtttttt	tttttgggtg	agacgggggt	ttgtgtgtgt	ggccggggctg
74281	gtctccagct	cttagccggc	agtgatccgc	cagctccgcg	ctccgggggt	gccgggattg
74341	cgagcggagt	ctcgttcaat	cagtgctctg	tggtgccacg	gctggaagtc	agtggtctga
74401	tctcggctcg	ctacagcctc	cacctccacg	ccgctgcctc	tgcccccaca	aagtgccagc
74461	attgcagcct	ctgcccaagc	gccacccctg	ctgggaagtg	aggagcgtct	ctctgtggcc
74521	accactcgtc	tggtatatga	ggagcctctc	tgctgtgctg	ccaggtctctg	aaagtagaag
74581	gcgtctctcg	ccggccgccca	tcccatctag	gaagcagagg	gcgctctctc	cccgccctct
74641	tcccatctag	gaagttagga	gcgtctctgc	ccggccgccc	atgctctgag	atgtggggag

74701 caccctctgcc ccgcgcgcct gtctgggagtg tgaggagcgc ctctgctggc gcgaaccttR  
74761 tctgggaggt gaggagcgct gcgcccgctc gagaagttag gaaacctctt gaacctctct  
74821 gccctgcacac ccgcccgctc gagaagttag gaggccctcc gtcggcgagc acccccgctt  
74881 gggaaagttag gaggctctcc gcgccgagc acccccgctc gggagggagg tggggggggg  
74941 cagcccccgc ccgcgcgcgc ccgccctacc gggaggttag gggctctctt gccgcggcgc  
75001 cctactctgg aagttaggag ccctctgcgc tggcagtgcc ccccgctcag gaggagagtg  
75061 ggggggttag ccccccgcgc ggccagcgc ccagtcgggg aggttagggg cgctctcgag  
75121 cgccgcgcct tactgggaag tgaggagccc ctctgccggc cagcgccccc ctctgcccgc  
75181 ggggaggggg gggtagcccc cctgcggcgc cagcgccccc gtcggggaag gagggtgtg  
75241 gggtagcccc ccgcgcgcgc cagcgccccc ctctggggg tgagggttgc ctctgcggg  
75301 ccgcctctacc tgggaagtga ggaagccccc tgcggcgcca gccgcggcgt ccgggaggga  
75361 ggtggggggg tcagcccccc gcgccggcgc ccgcccgtc ctctggggg agggaggttg  
75421 tgcggcgccc cccctactgg gagggttagc ccccccgcgc gccagcgccc cggagaggtg  
75481 gaggagaggt ggggggttag ccccccgcgc gccagcgccc cagcgaggtg gggagaggtg  
75541 ggggtagcac ccccccgcgc gccagcgccc cagtcgggga gggaggtgtg gggatcagcc  
75601 ccgcgcctag ccagcgcccc cgtccgggag gtgaggggcg cctctggcg gcccccctc  
75661 ctggggaagt agggagcccc ctgcggcgcc agcggccccc tccgggaggg aggtgggggg  
75721 tcagcccccc ccgccggcca gcgcggcgt ccgggagga agtggggggg gtcagcccccc  
75781 gccccagcca gcgcggcgt ccgggagga ggtgggggga tcagcccccc gactggccag  
75841 ccgcggcgt ccggaggtga ggggcgcctc tgcggcgccc cccctactgg gaagttagga  
75901 gccctctcgc cgtgctttaa ggcagcatgc tctgtaagag tcatcacac accttaact  
75961 taagtaccga gggacacaaa cactcgggaa ggcgcaggg tctctgctc tggaaaacca  
76021 gagacctttt ttactttgtt tatctgctga ccttccctcc actattgttc tatgacctg  
76081 ccaaatcccc ctctgcgaga aacacccaag aatgatcaat aaaaaaaa gaaacaccca  
76141 aaaaaataaa aaaaaataaa aaaaaataaa gaaagccgccc tgacctgtat acagattctt  
76201 gaaagaggtg tgcgaggtgc catgtccaac ctccgcgcgc gggggcgaga gaggagtgag  
76261 gccgagccgg ggccatagca gccgggttaa atggggtag gctctgtcca gacctctcca  
76321 cctcggttgc agccgcagcc tctccgctgc cggctctctg ctaacgcgcg gccacagtag  
76381 gcgcagatlt ctggcgacca gaccactgc agctcttctg tgccttggcg agcctgtctc  
76441 ccccgccagc agcacccctc cgcctccctt ttaccacggg cctcagcgcg gctgctcctg  
76501 gggctgcgcg cgcttggtgt taYtccagga cgttgggaaa gaacgggggtg gaaTgtgttg  
76561 ggtgggggtc aaagaggaaa cccagagatg caggggcgccc ctttcccggtg cgttgcggcc  
76621 aattgtctagc gcagggcagt cagggtaggg cgtctccctc ccaagtttat atttattttt  
76681 atttattatt tatttttttt accctcaagt ttattatttt ttattatttt atttattttt  
76741 gagacggagc ctctctctgt cgccagagct ggagtgcat ttgtagctca ggcacgatct tggctcactg  
76801 caacctccgc ctcccggttt caagcgattc ttctgctca gcttcccgag tagcagggat  
76861 tacaggttga tgccacaaca tccggctaatt tttgtattt ttagtataaa cagagtttca  
76921 ccatatttgc caggctggtc tcgaactcct gacctcaggt catctgcggc cettgcccct  
76981 ccaaagtctg gggattacaa gcatgagcca ctgcacctgg ctaacctcca agtttaaaga  
77041 cagccgcgag gccagtggc tcaactcgtt aaccccaaca actcaggagg ctgagggcag  
77101 gaattttaga ccagcctggg caacatagcg agaccccgcc tctaagaaaa attagggcag  
77161 cagcgttggg tacgtctgta atcccagcac ttggggaggg tggcgcaaaa gaggctgctg  
77221 aggggggaaa aatcaccctg ggggtagttg tgacacacta cagctctcgg tacttgtagt  
77281 gctgaggtgg gaggatcatt taagtgggag gctgcagtag gctactatgg agcgactgga  
77341 ttacagcttg agcaacagag cgagacccca ctcccaata aataaataaa gatagctccc  
77401 aaagatgtca cttgcttccac tttagcattt ttattgaaca tatttaggaa atataaacc  
77461 atggttaaag aaattgcaat tagtacaatat acacactaca cacatatgca cgtgcacaca  
77521 ctgaaaactg tcccttccag cgctcctgc ctggataatt tttttttttt ttgcaacgga  
77581 gttttgctct tgttgcccag gatggagagc agtgggcgga tatcggtcca ctgcaacctc  
77641 cctcctccgg gttogagcga ttctcctgct tcagcctcgc gaggtagctg gattacagcg  
77701 gccagccaca acaccgggct gatttttgta ttttttagtag agacgggggt ttgcccaggt  
77761 gggcagggct gctgggaact cctgagatcc gccacccctg tctctcaaa gctgctggat  
77821 tacagggcctg agcacttgcg cctggcctaa agtaattgtc ttctatttgg ttctctgccc  
77881 ttgggtctca ccaaacgcgc caactctaat ttcttaacta cccataggaca ctagcagact  
77941 gagagcccaa cagttaaaac ccttcaatag ctctccactc cccataggca tccagagct  
78001 tatctctgcg tctctgctgc cgcccccttc ctccactgag cccataggca tccagagct  
78061 cctcagagcc atctgggctg ttgtctgtcc tgccacacac ccttcccacg cgcctccctt  
78121 ccaagtctca ctgagggctg acctacacag ctgtctcagt ttctatttga ttctctggc  
78181 gctcctctgt aaagtctca tcacagccaa tgatactgaa actgttttct tatcagggcg  
78241 acggaaaagt ctattcaact cactgcocaa tctcctctcc ttggtcagaa aaagaccca  
78301 gtcagcagct tgcatattct gcagggcatga gtgaataata ataatgccta accctatat  
78361 agtgtcaatt ccacgcctgg cactgttcta ggcactcata taattcaca tatccacaa  
78421 aaccccaaaa ctgatgatct ctctctcatc taccagcagc tgagcagtta aataacttgc  
78481 tccagatatt aaggggtcga gctgggggtt gaagcctggt taccataaaa tgaaccaaga

78541	actggaagga	ggacaagaSc	ctccgagaagg	agtcaggtag	ggcggtatct	gtgcgcttta
78601	catctaaagt	cttccagctc	ccagggagacc	cggtttcatag	agcaggagat	agaaggtctgg
78661	aggagacagg	agagcctcga	gagccgtgtg	gagcaagcgg	tgccttttgg	gtccggggag
78721	caaagggcgt	gcctggatgc	gcggggcgccc	ggagcagcga	gtcctcaagc	caaactacaa
78781	ctccccagac	ccagcggggc	ctgcccgcca	ctgcagctga	ggcgccggga	ggggcgagcg
78841	ggctggggcgc	ctggcgagtg	gactgttcca	gccttctcgc	tggggaccgg	gcctggcttc
78901	cgcccccgcc	gtaagltggg	cgccccagc	ctactcagtc	cgggagagct	cgcgccgcga
78961	cgctcccgag	ctccatcaca	gcgcggggcg	gcagacgggg	gctgggggact	ccatattggg
79021	ggcatccggg	ccgaaccaag	tgaccgcgct	gggggggtccc	agcttcgctg	ccgtgcgcga
79081	ccctcccaag	ccggccccag	ggggcccaagg	ctggtgtcgc	ccgcgccctc	gcgcgctctc
79141	caggggccgg	gtttgaaggc	gctgggcagg	caggggcgag	aaaggtctcga	gagaagggta
79201	cccggagacc	cggggcgctg	ggggcgaggt	tctggggctgg	ggaaatagaag	ggggctctct
79261	cccgacagcg	cctccccacc	ccagtagagc	ctcgggttgg	cggttggagat	cccccgggag
79321	gctaggtccc	ttggggcggg	ctgtgtgcga	tctggggaga	atgtgtcggg	ttctgtggat
79381	gtctggggg	gtctggggag	accgatgcac	aggtggagag	ccgatctgct	ttccggggct
79441	tggatctctt	acaactctct	ctcccccgcc	ccggtagatg	cggaagaaMg	ccccaccaca
79501	gcactctcga	gcactctcga	cgaccctgg	ttcctgaagt	agtcctagc	gatcctcagt
79561	ccacaccccc	cttgcacctc	tttgggtcgc	ctgggtctct	ctggcgctcg	tgctgaggag
79621	ccctagcgccc	accgggtctg	aaaggagcaa	gacgatgatc	cccgagctgc	ctccggggcg
79681	cggtccgggg	ggggggcttc	cgctccggcc	cgctctggga	gctccgggag	gtcccaagac
79741	ccgctcgagc	tcgggccacc	acacacacca	cgtagagatg	gctccgggag	gtcccaagac
79801	cgctcaectc	ttttacaacc	agtcggccag	cgacccggca	gcggagaaag	gtcccaaggg
79861	ggagcccgac	ccagggtccg	gRatgtaggc	gggagggaga	gtgttggggg	ttctctgtct
79921	aaggcctctc	ttcctctcta	cgctcagtc	cgcttaacgc	ccacatagat	gctctcagat
79981	ggcgcgtctc	aggacggcag	ccacttctgt	gtaagattca	cgccctctct	tttctctgtg
80041	gatcctggag	ctctcccaag	cactcaggct	ccagccccgc	caattctctc	caatttctct
80101	cagaaaaagt	ctcgtgtacc	cgacgaagaa	cttccagtga	ggattgtctc	ccgatacaag
80161	ggcttctcgt	gccttctctt	catcattggc	tgcaacccca	ccatctcga	cggtgtaagg
80221	tagaagggac	cttaggtcag	cgggccaccc	tgccccgggg	gcaagttggg	agtcctgggg
80281	ccagatgtgc	agacgattgc	ttgcctaaag	ctgcagggc	cacacaggat	tcaaccacca
80341	gccttcagaa	gccaaagggt	tgattaccag	gagcctggaa	gggtcggaat	ggggggtttg
80401	tcacgtgggt	gaccagctgc	gtgggtatcc	ccatgggtgg	gtgggggtgg	ctgttctctg
80461	ctcagtgccc	atgcggcttt	gtgaattccc	acacctcttc	cttgcagact	gagctatata
80521	tcctgtgctt	ccagaagctg	acagacttcc	ctccggtag	tgctggggca	gagcagggtg
80581	aggggtcgag	aggttgggct	tggaaccacc	ttcctcatga	ctctgtgacc	tgacatcaa
80641	ggacccaggc	gacgagcccc	agtactgcca	gctggtcgca	cagctgtgac	atgaccacaa
80701	ggatgtggtg	accctcttgg	cagaggccct	acgtgagagc	cggaagcaca	tagaagttgg
80761	ggcagcaaaa	gagaggcccg	gcctcgttgg	ggtgggaagg	gcacgggatt	cttgagacctc
80821	actctttaca	ggatgaaaag	ctcgtccgct	actcttttga	caagacgctg	actctcgagc
80881	ttggaatccc	catgttggcc	acgcatcacc	tgccgctgca	tagagcaaa	gtggggctgg
80941	gggacctgag	accaccctgg	gaacattaa	tgagacagag	gagacatggc	tggggatccc
81001	gctcaaggcg	cttggggcgt	aggtcgttgg	gctgggtgct	tggggagact	cgcaagttgc
81061	cagcatcttg	gggtggggct	aggggcgtgg	gtagtctcga	ctctctttct	cgccgccagc
81121	tgaacttttc	ggcatcatct	gtactcgtct	ctcaccaaa	aaagattattg	aggaagttgc
81181	ggactttggc	aggtgagcca	agaatggctc	aggggggtgg	cagacatctg	ggggcaggaa
81241	gcttgggttc	tgagcccttg	ccgggggcat	gatctcgggg	gcagcagggt	tctcaaccat
81301	ggcatcttgc	acatttccag	ccagataatt	ctttgtcaYa	ggggctgcgc	cgctgcagtt
81361	aggaagtcca	cgacatcccc	tgccgcagc	agtagctgct	agttgtgaca	acacaaaagt
81421	ttctctgaca	ttgccatatg	ttacttaggg	gggcagaatt	gtttccagatt	gcaaacacct
81481	gttggggagg	ccctcgactg	aacctcgct	ccatccgca	gagcgtctgt	tgagacacaa
81541	tatggcaatg	cgccccgtgt	ccgcatcaat	ggccatgtgg	ctccccgtg	ccccctcttc
81601	ctcatgccac	tggactacat	cctgcgggag	ctgctcaaga	atgccatgag	gtgggggtgc
81661	ttgatgtgct	ggcttggggg	Yggacaggaa	ccgggNtgc	tgactacact	ggctcttccc
81721	ctctcgatag	agccacaatg	gagagtcacc	tagacactcc	ctacaatgtg	ccagatgtgg
81781	tcatcaccat	cgccaacaat	gatgtcgatc	tgatcatcag	ggttgcctgt	agtgaggagt
81841	gagctagggt	ggatgggatg	gggggtctagg	cactgtttct	ctgtgtattt	aggaccttga
81901	gcccccttgc	gccccattct	gggaacttgg	ccctgaccag	acaaactatg	ctctgaaatc
81961	tgagatggcc	atgagctgct	tatttaatgga	tctggggcca	gctgcagggc	ccagatgtgg
82021	gcctctgtca	gcagctgagg	agcttgaatt	tgagaaatag	tcagagtgct	gtcttatgct
82081	ctggggccag	gataaatgtc	acatcctgtg	agaaggtata	agcagctcag	ggccctggca
82141	ggggtaggag	tgataataac	aaggcccaag	ggctctaggtg	gaccacatct	cagctctggg
82201	tggaaaggac	aggaaggcag	actttgcact	gtctgtctgg	gggggttgga	tgcccccttc
82261	aaagctgagc	caagcccat	gttgttgcca	ctctgttagg	atctcagacc	gtgtggggag
82321	aatcgctcac	aaagatctgg	accgggtcat	ggactaccac	ttcactactg	ctgagggcac



82381	cacacaggagc	ccccggatca	gccccctctt	tggccatctg	gacatgcata	gtggcgccca
82441	gtcagactccc	atgcacgggt	gagaccctgc	caggccagga	tgaggggggt	agccacccca
82501	ggagactcaaa	gcctctgaag	cctcctgtcc	tgctcccctg	cccacccctca	gtcttgacct
82561	cggatctgcc	acgtcacggg	cctacgcgga	gtacctcggt	gggtctctgc	agcttgcagt
82621	cctcgagggc	attggcacgg	acgtctacct	gcggctccgc	ccacatcgag	gcggggagga
82681	aagctctccg	atctgaacct	acagcctttg	gcctgtctac	ccacagcagg	ctggcgccat
82741	tccctgcagg	acctccccgg	tcaggcaggg	cggcccctgt	ctccacacac	tgctgcactc
82801	tggtgtctcag	ggaccacagc	agatggactt	acatggagct	ggcactgtcc	gcctctcaac
82861	agggtccatt	gcctcctcgc	ctccagaact	tgaggcaggg	aagtgggcac	cttgaggcct
82921	ccagcacagg	tctcgtcaat	ctcgttctgt	ggaaaccccc	actctgaact	cccttataag
82981	ttcacattttt	gaatgccttc	ctcgggcccc	tggtgtggga	gggcaggtga	acctttgttt
83041	ctgcgcccat	tcaggttcaac	tgagcccttg	ggttgaactg	gtctgtgtcc	caagtctcta
83101	cctgcctcga	gagcctggca	ggccaggagt	agaatgggtc	ccaagtctgt	tgcatgtttg
83161	attttgttgg	agtgggatga	ctgcagcacc	ttatacaaa	agctttcatt	caatgttttg
83221	aaacaaatgt	tcgggggtccc	agataaattt	gaaggcccaa	actgaccacg	cttcgggcat
83281	cagtttttgc	tcttctcttc	ctggcagtc	cagtttctag	aggtgaagg	accagacgtg
83341	ggcaaacctc	tgagccaaat	gcttcccaag	cctgagttag	ttaaaataac	tggtctctgt
83401	ctgcgccagg	aaaagaacat	acaagtttgt	gccttggcag	gcctccagag	gttggtgggtg
83461	ccccactcga	aggaagagtg	gggcccctgt	agaaaggacc	aaggaatttg	gcgaagrtat
83521	caggttagct	caaggttaga	cctgaatcag	aactccagat	gcacatcttg	tgagggaact
83581	ctcacaccac	ttgcaggaga	agaaaggcct	aagggcgccc	tggtggggct	ggaggagaaa
83641	cttggaagtt	ctcttgccct	caacatgtag	ctccccagc	aaacttctgt	aggtcggtgc
83701	taggcctgtg	ccatctccta	cccttcacgg	ggatggaggg	gaagtgtgat	ctggaagcca
83761	aatggcagg	cctaggaaac	caacgtgact	tgctagactg	aaaatcccg	gagctgcga
83821	ggcagggtgc	tgaggtctga	gaggcaggca	gcagtacag	gccaggggccc	tgaaacatct
83881	gattttctct	gagccatagg	gatccatggg	tgagttttta	tttatttaga	caatgggtgt
83941	tgctctgtgt	cccaggctgg	aaataggtgg	ctgcagagtt	cactgcagcc	ttgaactcct
84001	gggatcaaga	gattctccca	ctcagccctt	ctgagtacgt	tggaccatca	tcaggagcta
84061	aatttttaaa	tttttttgtg	aaacagggtt	ttcataaagc	ccctatgttg	ccccgggtac
84121	acttgaaact	ctgggctcaa	atgatccttc	ccccagccc	tcccagaagc	gtgggggttg
84181	agggactgag	cactgcagct	ggcccatgag	tggtgtttga	gctgggaagc	gtgtttctgt
84241	gttggagctg	ctgagaggat	tcagtgtccac	gtgatttctt	agaaagatgc	tccagaaaca
84301	gagtagggga	agttaggaag	gggaagggag	gaagccaagc	aaggaagtga	cctcaggcaa
84361	agaccacaga	ccagtcattt	atgcctcagg	gttgaaggta	agagagctga	acctcagagt
84421	tactgatata	tttctccact	tggcagtcac	tggttaaaat	cagttgggaa	agtgaaacag
84481	tctattaaac	taaggatggt	tttttaagaa	gagcctcagg	tgctgggtgt	ggctttttaa
84541	agcacatcaa	aggtaatctg	ggcacacaga	aaacagcaag	actcccagag	gatctgggtg
84601	gagcacctca	attgtttttt	tggtttgttt	gtttctgttt	gtttttttaa	acgggacttc
84661	aatctgttgc	tcaggctgga	gtgcagtgcc	tggtcttctg	ctcaactcaa	cctccgcccc
84721	accccccccc	aaacccagg	tcaagcgatt	ctcctgcttc	agcctccgga	gtagctggga
84781	ttacaggcgc	gtgcaccac	accagctaaa	tttttctatt	tttagtagag	atgggggttc
84841	accacgttgc	ccaggctggt	ctcgaactcc	caacctcggt	atccatccac	ctcagcctcc
84901	caagtgtaga	ggattacagc	catgagccac	catgctctgc	ggagtgttct	tgatttttaa
84961	cctctgagga	accatccgc	aggccctgag	catctcaact	ctctcagaat	tgtttccaga
85021	cccaataaac	acattataaa	tcaaaacaga	ttcagagaat	agccaaaggg	aatgtttact
85081	gactacacac	ccggtctggc	actttgcaat	acacttgtat	attctagaag	cgatagtgct
85141	aaccgtttca	tagttatatg	attgatagtt	atacatgctt	aaactgctgg	gatgtgttcc
85201	aggaccggct	gtgaataccg	aaatctgcag	gcgctcaagt	ctcacagctg	gcctcccaaa
85261	acagcagata	tgaagtcaag	cttctagatc	tggtgggtct	gcactcctac	aatatttctc
85321	ttcctttctc	ttctttttcc	ttccttcttc	ctcttttttt	ttcgggtcca	ttttagagat
85381	agtcctgttg	tgctggccag	gttgaggtgc	agtgggcgga	ttcgggtcca	mtgcaacccc
85441	accctctctg	gttcaagcag	ttctcctgcc	tcagcctccc	aagtagcttg	gattacaggc
85501	acacgcacac	accctgaact	gttttgtatt	ttcagtagag	acgggggttc	accaagtgtg
85561	ccaaagctgt	tttgaactcc	tgacctcaag	taatccactc	gcttcggctc	cccaagagtc
85621	tggtgattca	ggtgtgagcc	accgcgcccc	gtcttttttt	tttttttttt	ggagcagagt
85681	ttcactcttg	ttgtcccaagc	tggaagtga	tggaacaatc	tcagctcaac	acaaactctg
85741	ctctccagagt	tcaagcggtt	ctcctgcctc	agcctccaga	gtagctggga	tcaaggcat
85801	gcggccacaa	cctctggcta	atttttgtatt	tttagtagag	atgggggttc	ttcatgtgtg
85861	tcaggctggt	ctcgaactcc	cgacctcagg	tgatctgcct	gcctcgccct	cccaagtgtg
85921	tggtgattca	ggagtgaagc	actgcgcccc	gcctcctttt	cttccccccc	tttttttttt
85981	agagcaggtc	tctgtcaccc	aagctggagt	gcagtgaggg	gatataagct	caactcagct
86041	cgacctcttg	ttgtttaaagc	atccctctgc	ctcagcctcc	tgagtagagt	gagtagcagg
86101	tgccggcccc	gagggccagc	taattttttt	tttccccaaa	atttttagta	gaagagaggt
86161	ctctatgctg	ccaggtctgg	tcttgaactc	ctgcctcgaa	gcgactccct	ctgttgagtt

86221	cctgaagctgc	gagattacag	gtgtgagcca	ccatacctca	acactgtatt	ttcaaccgcc
86281	cttcctgttc	atctccaaag	gtgggacatg	cggtatggga	gggcgcgattg	tRtatgggtg
86341	gagcatcac	atataaatgg	ctttaacctt	tactgaactc	caacagaccc	tcagtggaat
86401	ggcgtgatc	caagctcaatg	ctcccgaggt	ctcccgaggt	caacacattc	ctctgctcca
86461	ggcgtgatc	tagctgggac	tacaggggcc	cgccaccacg	cccgctcaat	tggtttgtat
86521	ttttttttt	tagagacgga	gtttcaactg	gttagccaga	atagctctga	ttctctgacc
86581	tcgtgatcca	cccgctcagg	ctcccaaaag	tgtctggagt	acagcgctga	accaccgcac
86641	cgccgctttt	tatttttttt	gagatggagt	ctggctcttg	gtcccgagct	tggaagtcaa
86701	cgccggatc	tcaggtcaat	gcaactggct	ctcccgaggt	tcagacgatt	ctctcgcttc
86761	agctcccgga	gtagctggga	ctacagggtc	gtgccaccac	ggccgcttaa	atttttgtat
86821	tttagtagag	cgagagtttc	accgtgttag	ccaggtatgt	ctcgtatctc	ggccgctctg
86881	ggctctcaaa	gtcgtgaagt	tacaggcggt	agccaccacg	cccgcccaaa	ctcgtctctt
86941	cttttagact	tatcctgtga	gggtgaatta	tggcctgttc	ctggaccacac	ccgtttctgt
87001	ttccccgaac	caactgtatc	ccaaataggg	gaagttagct	cttcaacctt	caaaaatggg
87061	gcaactgggt	ggcacgggtg	ctcacgcctg	taaccttagc	actttgggag	ggcgaggcgg
87121	cgcgatcac	tgaggtcaag	agttcgagac	caagctggcc	aacagggtga	agccctctct
87181	ctactaaaaa	tacaaaaaat	agccggcgct	ggtggcgccg	gattgttaat	ccagctattc
87241	aggaagcctg	ggcaggagaa	tcgcttgaac	ccgggaggcg	gaggtctcaq	tgagccgaga
87301	tcgcgtcaat	gcactccaag	ctgggcgaca	gagcgagact	ccatctcaaa	aaaaaaaRag
87361	gggggggggg	gggagcgac	attaggccag	cgcgagtgtg	cgctctcagg	gccaccggag
87421	ggccgctctg	gctctgccct	gttaaaagtg	gggacgcgct	cgctgaaat	tggcacatct
87481	tggtctccgc	aggtaccagg	gcttttaaaa	aggttcaact	ggcgggcgac	gctctctcaac
87541	gcctataact	ccagcaactt	gggaggttga	ggctgcgaga	tcacgaggtc	aagagatcaa
87601	gactataact	gacacatggt	tgaaaacccc	tctctactaa	aaataccaaa	atttagatcg
87661	cgtggtggcg	cgccgctgca	gtccacagta	cccgggaggcg	tgagggcagag	gaattcgctg
87721	aaccYggagg	cgagggttgc	agtgagccga	gatcgcgctc	ctgcactcca	ggcggtggca
87781	cagagcatga	ttccgtctca	aaataaataa	ataaataaata	aaataaataa	attaaaaaaa
87841	ggttgatcgc	agtcacgcga	tgaggttaagt	ctgtctctct	tcctcaccca	tgagctcgag
87901	ccacaggtcc	gcacggaaaa	ccaggagctc	caagactggc	gcacactcaa	catccaaacc
87961	ggcgggccgc	ggggagcgcc	acagaaagcca	atcggctcct	ctcatactca	agaccccgcct
88021	ctcaaaactt	gaggagcccg	ccccacgcac	tgctgacccc	gagccaaaaa	aactgcggtg
88081	cgctccgatt	tggtctcaat	tccttcccg	cgatggcgcc	acaggggagct	ggcggtcgcc
88141	ttgcggcggg	gaacttcagg	gtcggcgggg	agggcgagcc	cgggcccgcc	gagaaatcgq
88201	ccgctcagg	gaccgcccc	tcgccggggc	gcgtctctcc	cgccaccccg	ggcgcgccgc
88261	agcccggaat	cacggtggag	atcgggaaaa	cgtactctgt	ccggcgaccc	gatagacact
88321	ggcgtgagg	cggggccccg	ggctgggggc	ggggcgagag	tcaggggcag	gggtgtgggc
88381	ggggcgctgag	gacaggtcgt	cagtgaggcc	aagatccggg	ggctgggaggt	ggagaggagg
88441	agggggcggg	cttgaggaaa	gaaocggcgc	ttccggcgcc	tgagaaaacca	gccgggttgt
88501	gggaggtcgt	tgaccttgaa	ttatgcccag	cgacgcctac	aaacccaccg	ctcaggccct
88561	caccaggtat	gttcccatc	caacttccctg	ccagctctta	ggcttcatct	ctttttcttg
88621	ctaacgcgtc	ttctccaccc	ttcttctgtct	ctcgtctctc	ttttttccat	tgtccctgpc
88681	agccatccgc	agagagaaga	ccttccagaa	acaacaggct	tccttctcc	aaactctctgc
88741	tgctcttctc	cattttcaaa	attaacccca	aactccttag	ggcgcatctc	attttttgtg
88801	ctctctccag	ttctagccaa	caactaggaaa	ggcgattgcc	accagcagaac	acaactgtgc
88861	ctctgaagac	cacacgcctc	ttaccacctg	tgccctttgc	ttggaatgat	ttttctcccc
88921	tttctctctg	tttgcctgcc	tagctcctac	ttctctctat	agcttctcat	acttttgtgat
88981	gtcatcctgt	attccctctc	agggcaagtgt	agtggttccc	ttctctatgt	tcctgcaaca
89041	tttttttctc	aYctcagtc	tagtttttgt	aaacattatg	tgtaattttc	ttctgctatt
89101	ctctctccat	atactgggaa	cttctggagg	gcagcacttc	tttgtattca	tcactgtgtc
89161	ctcgtaccgc	ggcaacaa	cagcataggg	ccagacagct	agtgggtgcc	ttactcattt
89221	attgaatgct	ctgctaattg	tcagggtgcg	tgctaaaaat	tttgagggca	tttaactccc
89281	ttctaaaggt	gtctactatg	tcctttttgt	cttttttttt	tttgagagca	tttgagacag
89341	attcgcgtct	gtcaccacag	tgaggatggc	gtggcacgat	gacagctcag	tgaagcctag
89401	acotctctag	ctcaagtgt	cctccacact	caccctctca	gagtggtggg	attactggca
89461	tgagccactg	tgcccagctg	ttatcccttt	tcacagatga	agagactgag	gctcagaaag
89521	atggaataac	ttgtctcagtt	acacatagct	aggaagttag	gagctggaat	tttgtgtatt
89581	ttttttttgt	tggtttgttg	agatggaaat	gctctctgtt	gccaggctgc	agtgacagtg
89641	cgccatctct	gctcactgag	atctctgctc	cccgagatca	agcatctctc	ctgctcaaac
89701	ctcctgagta	gctgggaact	cagttgtgctg	ccaccacacc	cagctaaatt	ttgtattttt
89761	agtagagagc	aagtttcaac	atgttgggca	ggatggtttt	gacctcttaa	ctcgtgattt
89821	accaccaact	ggcctcccaa	agtgctggga	ttacagcgct	gagccacgat	ggccagacct
89881	ggaattatac	attaatgtag	gttatcttaa	tcagagacca	gcaactatga	tcctcgacca
89941	aatgaacaac	tgccgggtatg	tggtaaaaat	agcattgaaat	ttagagtacg	acttaaatct
90001	agtttaagta	tagctcttaY	cacttttcaa	atctgtgacc	ctaggttaagc	ctcagttttg

90061 tttctgttttt tggggttttt ttttttgaga cagagttcttg cttcttgcac ccagggccac  
90121 tgcctagttg acatctccgg ctcactgcga cctctgtccc ccgggtttcaa gaacttctcc  
90181 tgcctagttg ttgctctccc ggttcaaagca attttctctgc ctacagctccc Ygaattctctg  
90241 ggtttcaaagg tgtgtgacc catgctctggc tttttttttt ttcRctcttt tggtagagat  
90301 ggggtttttg catgttgacc aggtctgtctg caaactctctg gctctcaagtg taactgccat  
90361 ctgggctccc agaagtactg ggtttacagg tgtgagccac acaccagccg agttattttc  
90421 tcatctgtaa aatgagaata acagccYcaa cctgatccac ctacagtttg tggcagttgat  
90481 tgaagtgaat ggcagagtg gaccagtgag gtttagctgat gttgtggccc taactctctga  
90541 aactgtctct catgtggcag gatctctttt gttccagctc taactctctg atgacctag  
90601 Ncttttttoca Tcacagattc tgcaggaagt atocagttc gagtgaacga ccaggaggc  
90661 cgagaggaa tctatgtaca ctacgtggcg tgtgagttgc ttggcacatc tggcgctggg  
90721 tgcRggagt tggctgccc ggcagcactc ttatggcca tacttaacgt taacggcgg  
90781 ctggagcagt gggtagacaa gaaccggctg gcgtgacca agacagtgaa ggtgtgctga  
90841 cagaagaact cagaagaata cctgagcag ctgcagagc agctgagcg caagatcact  
90901 cgcaacccaa agcgaagca tggatgagatc aaccatgtgc agaaggtccg gatccctcc  
90961 ctaccagcg ccagaggagc ccagcttctc tgccagttcc ttgtggcttc tggcccaca  
91021 gtgccccaa catagcaaat cccatttctt aaactctctg agttgtgcac ggaactttacc  
91081 tacttctct tacttacttt tcataggtaa gaaaacagac agaggtgtaa gctgtctgt  
91141 cagctcttcaa gcaaggaat ggcacagctg ggaactagaac tcaggctccc ggaactggta  
91201 gctctttttg tctctagta gctctgtctt agtagaaaa atgactgaa gttcagttgg  
91261 aaagcagca atgtttattcg gcagttgcta tagccaaagc cctcgccaa gttctgtat  
91321 gtaataagct tccaaatctc tccaaacaa tccaacaca ggaattgta ttgtttgat  
91381 aggtgaggaa gcaggtctgag agagattaa taactgtccc aaagtcacac agccagataa  
91441 acggtgggct gaaattagaa taataatgcc aaacactgtg tgcccggtga cagaagatt  
91501 tcaactaaat ggttttttgt tgtgtgtgtg tgtgtgtgtg tttttttttt gttttttttt  
91561 tttttttttt agcaggagtc ttgctctgtc gccaggttg gagtacagtg aatgatctc  
91621 ggtctactgt aaactccgce acctgggttc aagcaatttt cctgtctcag cctccagtc  
91681 agctagact ccaggtacct gccaccagc ccggaataat ttgtatttt tagtgaag  
91741 gaggttctgc catattgttc aggtctgtct cgaactcctg acctcagtg atctgcccac  
91801 ctggcctctg caaagtgcgt ggaattacag cgtgagctac tgcgcccgc taagetaag  
91861 tgttttaata catgacctca ctgacattc acagtaggcc tgtgaagtaa caatgcagaa  
91921 gactctacga taagttttag gaaccagaca ggaatttgat caaccagca gttgtgcta  
91981 cctcagagcc agaggcctta gcaatgattt tggttttttt ctttaagagac gggagtcttg  
92041 gttgtgtgct caggctggag ttgaactcct gggctcaaat gttctctctg ctttagcctt  
92101 ctaagttagct gggactacag gagtgcattc ctgtactctg cttgtttttt tctgtattg  
92161 atacttaata attgtacata ttattgttat ttttttttt ttgacagcg tcttactata  
92221 tgccttagcg tgaagtgcag ggcacaaatc tggctcactt cagcctcaac ctctcggtc  
92281 caagcaatcc tctgcctca gccctctgag tagctgggac cataggcgag tggcaacctg  
92341 cccagctaat tttttatttt ttgttagaga cgggactctg tatgttgccc acactagctg  
92401 caaactcctg gctcgaagc atcctctctg cttagcctcc caaagtgcag ggttacagc  
92461 tgtgagccac tgcgcccagc caagtctcag tttctaaatt tataaaaaat agggagtac  
92521 accagagcaa gatttgtcaa agtctccaga caccagcaYa ttgttactt tgaactgta  
92581 cagctctcat cttctgagta catgagttag tggcgaagcc caactattaa attctcttaa  
92641 atggaagaaa gaaagtggat aagatagaaa aatcaggggt ttgcacttta ttaagtgtga  
92701 gtaattgttt gagaagcttt tatctcacac ccaataaag ttgtcagat taactctgtt  
92761 cagcctctac ctctccgat atctctagca atacttgtgt gatatcacg cctgtgtaca  
92821 gtgccaagt tacttctctac ttgtggtgta tttaaaaaat tgaatatccc taactgtata  
92881 gaactctca aggtctaaac tgcgaatcg aacttctat ttgaatggaa atttctata  
92941 gctgtactat ctatgacctg agccactagc cacatgtagc ttgtgagcac ttgaaatgct  
93001 tctgtgtga cgaagaatta agaattttag ctgggcccgg ttgcttaccg ctgtatccc  
93061 agcacttttg gaggccaagt tgggtggatc accaggtcag aagatcgaga ccatcccgg  
93121 taacatgttg aaacccatc tctactgaaa aatacaaaa atatagccgg gcgtggtggc  
93181 gctcactgt agtccagct actcgggag ctgaggtag aagatggcgt gaacccggga  
93241 ggtgagctt gtagtgagcc gagatcctg gctcactoca gcctggctga cagagtga  
93301 ctccgtctca aaaaaaaat aatttttaa ttatgggtg agtggtctac gctctaatc  
93361 caaacacttt gggaggcgga ggtgggtgga tcaactaagg tcaagtgga cagactgga  
93421 tggccaactt ggcacaaacc caatctctac taacttagga ggtgagga gggcgtggtg  
93481 gtcacacat gtaatcccg ctacttagga ggtgagga ggaagatcgc ttgagccgga  
93541 gagatggag ttgagtcag caaagatcat gccactgac tccatccagc ctgggtgtcc  
93601 gactcaagct gctcaaaaa aaaaattaat atcattttt aaaaaattt tatttttat  
93661 tttgtgagac agagtctcac tctgttaacc aggtctggat gcaagtggtc gatctcagc  
93721 caactgaacc tccgctccc agattcaagt gagtttctc ccgacagtc ccaagttag  
93781 gggactacag gcacgcacca tcatgctggt ctaatttttg tatttttgt agagatagg  
93841 tttcaactatg ttggcctca aactcctgac ctcactgat ccgctgctt ccgctccca

```

93901 aagtgctggg attacagggt tgagccaccg tgcccagcca aaatttttcat tttttgagac
93961 agagtctcac tctgtcacct agcgaggagt acagtggcga gctcttggtt tactgcaaac
94021 tccgctctcc tggcctaaagc attttttgtg cctcagctct ccaagtatcat gagattacag
94081 gcgtgtgcca cccacaccagc ctatttttta ttctttattt ttgtagaga tggggtttcg
94141 ccatgttgag cagactgggtc gcttgagcca ccagcttttg ctgagtataa cttcggtctc
94201 tcaaaagtgt gggagtacag cgttgagcca caggtttgat tcaagattaa ttctctcaat
94261 cccatactga caagttggcta ctctatttga cagtgtagat agcaggggtg atcaacggcc
94321 gtttttgcact caacaagagc ataccttga gt tggcaacc agcaggggtg atcaacggcc
94381 agtggtgact aaatcaaga ataggtaatg aaacaatcat caggttaaca atcaacggcc
94441 ttcttcagag cctaataaat gttaattctt aaataaattg caacataatg gaaaagtgt
94501 gtttttgttg ttctttatgt cttttgcaga agagtactgt attttgtgag aattgggaaa
94561 tgrgttttaa agagatgtgt gaggaggaaa atctttcttc ctaactctta caggtgttag
94621 gtccagaatt ggcagagcca ggggccatg cccaactctg ctgagtacca gtctgttgac
94681 cttgggcaag tgacctcacc tctttgtgcc tcagttttct ctaataaac agggacaatg
94741 agcactctta ttccaaaggc tcatgttgag gaacaaaatt aaataatgtc tcttgggtgc
94801 agtgactcat gctgtaatc ccagcactct gggaaagtga gatgtatgga gattgtgagg
94861 tcaggagtct gagatcagcc tggccaatat ggtgaagacc cgtctctact aaaaaataca
94921 aaaaaggcca gcgcaggca gaccacgagg tcaggagtgt gaggccagca tggccaacat
94981 ggtgaaccac tgtctttact aaagatacaa aaaaattagt ggtgttgatg gtgcacacct
95041 gtaatccagc cactttggga ggctggggca ggagaatcgc ttgaagctgc gaggcagagt
95101 tgcagtgacc tgagattttg ccatgtgact ccagcctggg cgacaggcgc aggcctccatt
95161 tccaaaaaaa aaaaaaRaana aaggctatgt ctcaYgcctg taactccagc ttcttgagag
95221 gccaaagggt gtggatcatc tgaaggtcag aaatttcaaga ccagcctggc caaactgggt
95281 aaacccctac tctaccaaaa atacaaaaaa tttagccgggc ttgctgttgg gcgcctgtaa
95341 tcccagctgc ttgggacact gaggtgggaa gatcgtctga actcaggagg ttgagtttgt
95401 aagtgagcca agttgcacca tgcctccagc ctgggcaaca agagcaaaat tcaactcaca
95461 aaaaacaaat aaattagctg ggaatggtag caggccacat tagttctcagc cacttgggag
95521 gctgacgtag gagaatcact taagcccaga aggcagaggt tgtgttgatg tggactctcc
95581 ccaactgcact ccagcctggg cgacagagtg agagtctatc tcagaagaaa aaaaaaaaaa
95641 catgtctgtg ccgtgcttgc cacttggaaat aaatggcagc tcaactagtc cgcagggtaa
95701 gggagctcac cttgccctgc ctgtttacta ctctgtaaaa ttgaaataac gcagctgggt
95761 gcagtggctc acgcctgtaa tctcagcact ttgggagtgc gaggctgggt gatcacctga
95821 ggtcaggagt ttgagaacag cctggccaac atggcgaaac tccattgcta ctaaacatac
95881 aaaaattacc ttgagtgtgc gcaatgtgcc tgtactccca gctacttggg aggcgtgagc
95941 agggagaattg cttgaacctg ggaggcggag gtgtgcagta gccagagtgg tgccactgta
96001 cctcagcctg agcaacagag tgagactccg tctcaaaaaa aaaaaaaaaa aaaaaagaaa
96061 taacactctta tgaagggtt actgaaagga ttaaatgagg ccatataaac atataggcac
96121 agtgctttgt ggagggggtt ctgggttaag cactggctcc tctctccctt ctactgtaca
96181 gggcagggcc ttgtctgccc tgggttaggt catggaccgt cttgggtccc ccttccattg
96241 tcaggctgta ctgggttctc cccacagggt ctggcctagg agcaggtagt gttgtgagtg
96301 ttggtgtgact cttagtctct gctttctaa gacttgaaaa atagagaaac ccaggtgggc
96361 atacttggac agtgttgcce tgttgaccac ttctcgtga acttgggaa gagggaatgt
96421 ttcttggagt ggaagaatcc agtcaggcct cctggagggg gtggtgaaga tctcaagctt
96481 tgaagtagtg gaaagagcac aataacaaca gtgacaaagg atagatcaca agacatcatc
96541 gactcacctg actcatgctc cgtgccaggg tgaggaaagt tttctcgtg gcagtgagag
96601 aaggaaaggc attctccagg tagaaggagc tgtttgtgca gaagcatgaa ggcgcctaaa
96661 agggggctgc tgatagcaat gaggaggctg gggctggagg cagggaagtc gaggggcct
96721 agtaaatctg ccaagcctga gatggggcca tgactacctg gagcagcatc gagagtgaa
96781 ttccagagag gccccaattc cagaagtgtc agaagttagc tYgacatcca ctgtattagg
96841 acttggccag gccaggggag agagagtagt ctgagccaag cctagtattg ggtcttgga
96901 actgggacca aggtgtgagt tcccagtaaa atgagcaacc caggaggagg cgcaggtttg
96961 caggaggagg tgatgaggtc cttctgcac atctttagtc cgtgtccctt gttggaggga
97021 cctctggggc agtcagatg gcaaggcagg

```

chrom 4 genomic sequence (SEQ ID NO: 2)

>4:36870401-36969350

```

1 taaaaaagca tttccataaa tatatacaat cctgtcttgt caatcaaaaa aaaaaatgaa
61 aaaaataatc caactaaaa taagatatata acataatgtg ctaaaataga caaagctgaa
121 tatctatct caactaatgtt attgtcaacc atgttcaatt gtaaaataga ctactactgt
181 ataattcttc ataacataag aagttgatgc Mttacaaaaa acttactctg gatagtttat

```

241	tgacttgact	gggtcttaca	aatatagatt	aggatatagg	atagaaaaagg	atttgactgt
301	aaaaattgca	gttaacagtg	ctcaaaagcat	agtgatactt	agtaattgatg	gaatttttgg
361	ctatgggata	gttataaact	acttctggca	atgtgagctg	actttatttgg	tgttaaaatg
421	tatcaatcat	cttataaata	aataaacagag	ctataataaa	caatttttgaa	caattttccc
481	tacttgccct	tttaaaaatc	agctccataa	gagacaaact	aatttaattatg	tgtttccact
541	attttttttc	ttacattaca	ctttgtctgt	gacttttaata	aggcagagaaa	atgttttgagg
601	gtgggcactg	ttattggggg	aggcgctaca	gtgaagactg	tgcctattct	ttggcgaggc
661	gattacatac	gtccaaacaa	tcccctggag	atcagtatgt	taagccctgt	tttttagtgt
721	atctcatcat	gtccaaaggg	catagataac	taaacctagg	gtccaccctc	agtcctggag
781	tagagttaga	gaaagggctt	ccaagagtg	aaccattcag	agttgtctctg	gagattaac
841	tgtagaagaa	gtgtctctgg	ttcatagctg	atgagctccct	tctatgtctaa	ctctacaaat
901	tcaggaggga	aacaaaaaaa	caaaacagct	ctttttccctg	ttaaagctcaa	tagcttttca
961	gattcccaac	ccatagtttag	acttgggaag	cattgttccac	tgcctggggg	gagggtacta
1021	tcaattggtac	tactatcaat	ggtagggagg	ctggnaattca	gaacccaaact	tcagatccag
1081	cagatgaatc	gttgaacaaa	ttcagaaaag	tgaattcttga	acctcaaaata	gtgagggctc
1141	caactaacac	atcaagcttta	agccacaaga	aatatctccag	cccccaacct	ttgatagacc
1201	tgagcctacc	tacctgtaaa	agccaggtct	gggttccacag	tgcctattttt	ctaccactgg
1261	tactctgtctt	gtatcacatt	gtgttttaatt	tttatgtctga	ctagctccatc	ccccctacca
1321	gacaatatcat	aaagatgggg	attgagctctt	ctcaccattt	ggtagagaga	gtaccattttc
1381	ctaccatgct	accttgccttg	tatcacattg	tgttttaatt	ttatgtctgac	tgttccattct
1441	ccccctaccag	acagtcacata	aagatggggg	ttgagctctc	tcacactttg	tattctccac
1501	attttgtaga	atgtctgggt	aattgtagcac	tcactatgag	tttgttgtaag	gacaaaaataa
1561	acatttgatag	ctttggccaa	aataaactcg	aaatttttaa	ggatatagaac	ttcaaatatt
1621	aagatctttac	actacagattt	tgcaggatgt	agcaggcaaa	agtgactagat	gctatgtttc
1681	agttattattt	acactacaacc	cagaccatga	gaaagcagtc	ttagcttata	tcattgtacac
1741	atagatatagt	ataggataat	aatgacctgt	acctatattt	agctctttat	gctgtacatc
1801	aaatctctac	ctattttgtt	ctgtgtcatc	tgaatgaata	cacagctaga	tggatagata
1861	gatgatagat	agatagatag	atagatagat	agatagatag	atagatagag	gcatgaggac
1921	caattttgaaa	atataagcac	tttatgagcc	acatgttttt	acattgtaat	catttggttt
1981	ctacaaaaaa	ttgaccccgag	agagggttaag	caacaacaca	tatcatactc	tttgttccag
2041	caatttagaggt	gctgtgggttg	ttaagatatc	aattataaaa	tattttatata	aaaacacaaa
2101	actttttttt	cttttgacaaa	gatgatgtgat	tattcacaaa	tagttttgtca	caactttgtc
2161	ataaaaaatg	gcactgtttc	caattcatca	aaattatagg	acttccattac	actaaaaaca
2221	ataaatcaca	catctaacag	caaacccgacc	cttaataacc	tttgttacct	acttttcaac
2281	atttcccgaga	gtgtgacatt	gaactctcat	gaggtaaagat	gctctatttt	ctagctgtct
2341	gtggcaaaaat	ataacagagt	aacattaaat	tctaattgtg	attttttgaga	tgaagggtga
2401	gtggtaattt	ttgcaattat	cttttgaaag	gtaagcaaaa	tatacttttc	tccttagaaaa
2461	agatttccatt	tctatatcta	cattataaac	catatatcaa	gtgatattac	gagaaaaaaa
2521	gtaagatcaa	gattacatat	taaaaaaaga	gaaatggggg	tgaagctctta	gactttgacc
2581	ctggaaagat	gatgaacagag	gagagagctc	ttttttggag	acttctccctc	ctttcaatgt
2641	ccatccccca	gcttctagag	cttctcccca	ctctggggag	attcatgtta	ctctccagac
2701	cccgagctcaa	tctcatgcat	caatgtcaag	taactccacc	tgcctaatct	attataaact
2761	tgaatttcaga	ttgaaaagata	agacaactgt	tcagatgac	tactttcttt	ttttttcttt
2821	cgccggtggc	gggggggctg	ccatatttaa	ctgaaagtct	ctttaaattt	agaaagtgt
2881	ctatagtttt	ttttttgaaa	tttttatttt	tttaatttat	tttatttttt	attatatatta
2941	tacttaagt	ttagggtaca	tgcgcacaa	gtgcaggtta	gttacaatgt	tatacatgtg
3001	ccatgtcgtg	gtgtgcacac	catttagcat	taggttatct	tctcaatgtc	gtccctcccc
3061	ctctccccca	ccccacaaca	gtccccagag	tgtgatgttc	cccttccctg	tgcattgtgt
3121	tctcattgtt	caattcccat	ctatgagtga	gaacatgtgt	tgttttggtt	ttttctcttg
3181	agatgatctta	ctgagaatga	tgttttccaa	ttctatccat	gtccctacaa	aggatattgaa
3241	ctcatcattt	ttatggctgc	atagtatccc	atggtgtata	tgtgcacaa	tttcttaatc
3301	cagtctatca	ttgttggaac	tttgggttgg	ttccaaagct	ttgtatctt	ttttgagagc
3361	gcaataaaca	tacgtgtgca	tgtgtcttta	tagcaaacat	atttatagtc	ttttgggtat
3421	atacccagata	atgggatggc	tggtgcacaa	gggtatttcta	gtctcagatc	ctcagggaat
3481	cgccacactgt	acttccacaa	gggttgaact	agttttacag	ccacacacaa	gtgtaaaagt
3541	gttctctatt	ctccacatcc	tctccagcac	ctgtgttttc	ctgaactttc	aatgatggc
3601	atttcaatgt	gtgtgagatg	gtatctcatt	gtggttttga	tttgcatttc	tctgatggcc
3661	agtgatagtg	agcatttttt	catgtatctt	ttggctgcac	aaagtctctc	ttttgagagc
3721	tgtctgttca	tatcctttgc	ccacttttgc	atgggggtgt	tgttttgttt	ctgttaaaat
3781	tgttttagtg	catcttagat	tctggctatt	agccctttgt	ccagtagagta	cttgcgaaaa
3841	ttttctccca	ttttgtaggt	tgcctgttca	ctctgagggg	agttttcttt	gctatgcaga
3901	agctctttga	tttaattaca	tcccatttgt	caatttttgc	ttttgttgc	attgtctttg
3961	gtgtttttag	catgaagtcc	tgtcccatgc	ctatgtctctg	catgttatgt	ccatggtttt
4021	cttctaggtg	tttttagggt	ttacgttttaa	gtcttttaact	catctgaaat	taattttttt

4081 ataaggtgtga aggaaggggat ccagtttccag ttttctacat atggctagacc agttttccca  
4141 gccacctttaa ttaaatagggt ggccttctcc ccaattgctt gtttctccag gtttctccaa  
4201 taccagatgat tttagatgat agtccggttat ttctgaggtg tctgtctctg tccattgatc  
4261 tttagctcttg ttttggtacc agtaccatggt ttgttctgtt actgtagact tgaagtatag  
4321 tttagaagcga ggttagcatga tgccttccag ttgttctttt ttgttagagt tgaacttggt  
4381 atgtgggctc ttttltggtt ccatatgaac tttaaagtag ttttttccaa attctgtgaa  
4441 ggttagcttga ggttagcttga tggggatggc attgaactta taaattacct tgggcagtat  
4501 ggtcaattttc acaataatga ttcttctac ccatgagcat ggaattgtct tccatttgtt  
4561 tgtgtccctc ttattttccat tgagcagtggt ttgttagtct tccctgaaga ggtcctccac  
4621 gtcccttgta agctggattc ctaagtattc tatctctttt gacgaattg ttgagggag  
4681 ttcaactcatg atttggctct ctgtttgtct gtatttggtg tataaagaatg cttgtgattt  
4741 ttgtcaactg atttgtatc ctgagacttt gctgaagttg cttatcagct taagttagatt  
4801 ttgggctgag acaatggggt ttcttagata tacaactcat ccaatctgaa aaggggacaa  
4861 ttgtactctc ttcttctcta attgaalacc ctttatttcc ttctctgccc taactgtgct  
4921 ggccagaactg tccaacacta tgttgaacct accagaatct ctgggacaca ttcaagcag  
4981 tgtgtagagg gaaatttata gcaactaaatg cccacaagag aaagcaggaa atctacaaa  
5041 ttgacacctc acaatccaaa ttaaaagaa tagaanaagc agagcaaaac cattcaaaag  
5101 ctagcagaag gcaagaataa actaaaaatc gagcagaact gaagccaata gagcaaaaaa  
5161 aacccttcaa aaaaataggt aatccaggat ctgttttttt gaaaggatca acaaaattga  
5221 tagactctga gcaagactaa taagaagaaa aagagagaag aatcaaatag accgaataag  
5281 aaatgatataa ggggatataca ccaccaatcc cacagaaata caaaactacca tcagaaata  
5341 ctacaacaaa cctcagcta ataaactaga aaactagaa aaactagaa atactccaa  
5401 cacataactc ctccaagacg taaccaggga aaactgtgaa tctctgaata gaccaataag  
5461 aggatctgaa atttggcga taactcaacg ctaccaacc ctacttaag tccctcagct  
5521 tggattccaa gccgaattct accagaggta taaggaggag ctgttacctt agcttctgaa  
5581 actattccaa tcaatagaaa aagagggaa ctcccttaac ctcttttatg tcccgagctc  
5641 catctgtata ccaagcctg gcagagacac aaccaaaaaa gagaatttta gaccaatc  
5701 cttgatgaac attgatgcaa aaactctcaa taaaactg gaaagccaaa atctacaga  
5761 catcaaaaag cttatgcacc atgatcaagt gggcttctac cctgggagtc aaggctggtt  
5821 caatataatg aaalcaataa atgataatca gcatataaac agaaaccaaa acaaaacaa  
5881 tatgattatc tcaatagatg cagaaaaggg ctttgacaaa attcaacaa cgttcatgtc  
5941 aagaactttc aataaattag gtgttgatgg gacgtatctc aaaaataaac cagctatcta  
6001 tgacaacccc acagccaata tcaatctgaa taggcaaaaa ctggaagcatc tccctttgaa  
6061 aactggcaca agcagggat gccctctctc accactccag catgttcttc actcaaatgt  
6121 cccataaact gctaactatg aagtaccaaa taataaatat catgtttctc atttggcttc  
6181 ttatattcaa taaaataggt cattatatta catatgcat caaatttggt aaaaataaca  
6241 ctttcccccga tgggactcag tgtaactatt tgcattctga gtgtggaaaa acctgtgtcc  
6301 catgaaaaag cttgtattct ttaagtggct ctactccaa tgtagataga tatggcata  
6361 ggtgcagcat tagatgcagg tataggtata ggtgttatt actgtacatc atgttatata  
6421 aggaattctc ttctccctac tcttttctg tctcttccaa aacaaaagct caggctatct  
6481 gcaactcagct tgtatgtatc agaaaacagt tagtaattta ttttaaaagt aacttcagga  
6541 agaaatggat gttgtgaatg aaggaattta aagcagcatg aagaataaaa cagtcattc  
6601 agtctgcagc gcagtctgtg gcagctattt ggggtataag gcaactatct tgcgttttaa  
6661 accaaaacccc attgccatg cagcaccatt ggagatttat gtaagtgttc aaccagtgcc  
6721 aaaaattcat cattcttagc tgtctactca ctataggtag ggccgatccc caaattctaa  
6781 aagattgaaa tggtaggagt cttgatacca ttttcttgt attgtttaaa atgcaagcc  
6841 actcaactac aaatttgaac agggatgtgc atttgccaaa catgtagtta taatttgat  
6901 ttctactgtg taaccttttg cgaaaaatga gatggattag catlcaaggt cactgaggtc  
6961 gtgttccatg gatattgttt cagtcacgcc aaccagggac agattgtgta cttattatct  
7021 tgtttcaag tattattcaa acatttaaaa aatatcccag cactatacaa catatagat  
7081 taatgtatgt gtaattttt ggtaatttt cgttgacttt aaatgatttt gaggttccaa  
7141 tccagatgtc taaaagatg gcacctagat tgaataactg aaagctattt ggtttaact  
7201 gcatgctag cttggcRlcc agggaaactcc tgaagaagcc agaaattagt atagcctcag  
7261 ttctaatata attccaaagaa gatttagcct ggtaacaaga ttgttctact actacaatt  
7321 caaaaagctt tagaaaaagc aatggaggac tacagcttag agaaaatita ctggtcagat  
7381 aaatcacaaa attatttccac tcttccctac tagactagta aatttttaaa actcatgtt  
7441 ttatattctc gtattttcca caaccagtt tttaaaattt attagacctt aatacaaat  
7501 tgtaaaattc atgtcacata tctgtgtgtg gtcttttaat tgaacctgtt taaccaacga  
7561 gttttaaatc ctgacagtg tagacctgga acgaataacc aatatgcatg ttattttcag  
7621 gagtgtgttg aggcattcca agtggcataa aaaaataatg tgcataagct gttaaacgtg  
7681 gaagtttatt attagaggt aagttaattc aattcaataa gtaatttaat tcaatgcact  
7741 ttgtagacat tggagggaaa aacagtcaaa aaattttact ctggaattct attctagc  
7801 aaggggagcc accaataaat atagatatg ataaacaatt gtgtgggtgt ttgaaggtg  
7861 acaaatgttt tggaaacaaa gaaaagtag agggMaaagg gtaatcttga atagtagtg

7921 tgtgggctgt tgccaaagttg cagtttttaa aatgatcaaa gtctgcctca ttggcaagggt  
7981 gagatttttaa caaaggtctt aggaagttag agactgtggc agagcaaggtg ttggacaaggt  
8041 gctctccagg caagggcaaac acctagagca aaggtccaaa ggtgagagcgc tgcctgatgt  
8101 gtccaaagaa gaocaggagac ataatgtagt cctccaggg tagtgcacgc aggggaagcgc  
8161 tgtagaatat aaggtcaagag agttaatgag gaagcacact cagctacata cagcctgaca  
8221 catggtgatt atatgaactt tagctctctg tgaatcagg gaagacacat tgcagaattc  
8281 tgagcagaga aacagctctac tgaagtgtaa gtttccaaatg gatgatctta tttttgtata  
8341 gagtacagac tatagagaga agaggaataa tgcagctgac caattgtgcag agggttgcag  
8401 taatccaggg acaacagatg ggggtccaga tcaagatag aggtgtctaa aagtatgtaa  
8461 ttgttgata ttgtttgaat gttagagaaa aagaacttct ttgtgatttt aataaggaaa  
8521 acatgtgaaa ggggggaaat caggtgact ttggcctgt caactcaaaa gatggaaatt  
8581 cctcaactga aatgagagaaa actgagatga gaagtctca agcagataga tggtaactca  
8641 gtctaggata cgttcaagca gaattgttag acagtgtgt atatgaattc gaagttcaag  
8701 tacaaggctc tggagtggag gtttaattt gggagtacc agcagataga tggtaactca  
8761 agcccaactgc ctggatgatt ccaacaagag agtctatgt tggaccaagg actgagctt  
8821 ggcaatacaa ctttaagtca ggacgaagag ggagattaac cggaagaagt cggccagtca  
8881 gttagaagaa aagcaaaagg ttgtgtgtgt tttttccaa tagctaagaga agaaagtga  
8941 tcaaaaggga aggagaaatc aatttggta aacactata aggtccaag gatgaggt  
9001 gaaaaaaaaa tggccgttgg gtttagtgac acagaactag ttcatttgta aatttaata  
9061 ttcaactctc catcagtaac gccactttct gaacactgac attttgattt tatgtaggaa  
9121 tttcccccaa ccaagtttta tgcctttggc tttctttat ttttcttat actgtctctt  
9181 tttttgtgt tttatgggtg aaatttggtt tttctttaa ttttttaact ccttataact  
9241 tatgtgcttt ctgcggacat gcttttgaat tacattttta aatttttaatt tcatttcaaa  
9301 tttataatat agcgaatgt ttgaattcag cttaactct ttatccactt ctatgtctat  
9361 attgattttt tggttgtctt tccaaatcac attcatgct tatttttagct gccacattgt  
9421 ataaagtctc ctagcattgt ttaactttca cacttagtaa gtatgatgt gtatgcattg  
9481 taactgtaat tgacaataaa catctgtttt atttccccat tttttgttaa ttcaacaaca  
9541 cagttgtatt ctactgtgtg caltcaattga acttgtctaa acttgcata aactcaccc  
9601 aatattttgt tttaactctc tctcttcaa cagcctcatg attatagca tttctcata  
9661 cttttatttg tattgcaaat tactgtaact taacttttgt tacttttatt tctggagatt  
9721 cttttgttac tgtttaccaa ctgtgacttt cattctggct gacctttgtta atttcttatt  
9781 tattcagatt ttgttcatct tgcatttgta ggtacattta tacaactcgg aatggcatt  
9841 cagtttaact agtgttctca agccaataaa tgcattgtcc tgaatgtgaa atataactat  
9901 taatataaaa tgagtgtgca tttagagact caatcagtag cagaatagag gtaataataa  
9961 ttgatagtta ggaattatga aaaggaggtt aaaggggaa aattttatag tgaataataa  
10021 ataaactctc agaaaggcca atgcaaaagc ttctatcta tgcgaagcca gagacata  
10081 catctaagtc atagctctac acaattgtca aatcctccc atttaaatag taataactat  
10141 cccaacatac atacacttaa acatcattaa acaagaataa gccaaggagc atattaatag  
10201 tcatctctct ctctctctct aatacttgct ctgcttata ggacagacat tttagttct  
10261 gacagacatc ttatttagat aggttcaatt agcaagttat gcaaggttag gctgttctc  
10321 tgaaaaaaaa atgaatgtca ttgaacactt cactgcacag gatgtcttca gtttttatt  
10381 tcccaataaa tagtaagcta ttagtagac tcttatgta ataattctg taattctata  
10441 tgtcaacaaa agaaaaatat ccttcagaaat taacttatgc ttcagagatc ttctaaagga  
10501 tttaaaaaaa tttagtagc cactctttaa aaagaaaaac actcaacaa ttaagaaact  
10561 tacatatgta attctctgtg actcaaaact caaaagtgtt ctttccacgg tttaacatt  
10621 gaagtgatga agaaaaacca gctattttat gttgatgttg ctttcagtta tttaacatt  
10681 cttaatgtat tttctggagg tctacaatga ccttagatcc catgaaatga taatataat  
10741 gtgaagtgtg gtcaagtact tgtataagat actccaatgt actccaatgt taacaaatt  
10801 tctctgaagt tatcaaggca gaagaatgca gcgatcatt agtccacgga ttaaatatt  
10861 cttcatagct ggttcaacta gctgattgat ccaatgagtg ttgctctgac tccccatt  
10921 ccatctaccc ttactctgag agtcacttcc tgatttccct actcggaag atgcctcaac  
10981 tctgaatttt ataattgtta tgttttatgt tgcagatgga tgaatatca ttttatttcc  
11041 aaagaagtgt gaattatttt ttttatctag agtctttgtt caggcaatcc ctccaccga  
11101 gtcccaaaaa ttgggaaaaaa gacttagctt ctatatcaat tttttagtct tttatttcc  
11161 tgtgttactt ttaacttgatt tccagctagt aacctaaaaa ccttacaaca tttagttctt  
11221 tcaagatcaa aaaaacaaaa tgcacagaca taactgaat ggagctgaa tttagttctt  
11281 tgggtgcgca taacaaatta ctacaactgc ggaggttaa aagaacagaa atttatctc  
11341 tcaaatctt taactctaca aatcaaaaat gaaggtgtca gagggtctc gcttccccgt  
11401 aaggctcatg agaaggtatc tttctgtctt ttccagctt cctcttttcc tataaagatt  
11461 ccagctcatg gattgaaggt ccaactagat ccactcaaat tctgaggttc tttagttctt  
11521 attacatctc caaagacact acttcaaaat aaggtcacat tctgaggttc tttagttctt  
11581 tgaatttttg ggaacatta tcaatccac tatagagata aataacataa atttatctc  
11641 gactgtgtg actctcagga cctcataga tctctttgga ttgtattttt ctttctgtct  
11701 aaaaattttg aattgggcca acaagggaa tctttaaggt tttattctat tttctgtctt

11761	atgaatctcta	ggatcaaagg	acaatgtcat	ccaagtgttc	aggettatgag	cacaaaataat
11821	tggcgttttc	tgtctataca	tgtacttacc	tacaaggtgc	acgatattatg	agtttggctt
11881	tgtaggacat	tttttcattc	tggagacatg	aaagatttct	tatcaaatgt	agctgaattt
11941	gggcaatgtg	tgtactcctga	ccatttttca	gtgatgttct	tggccatgat	attcagaaga
12001	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtatcccaag	tagctggaa
12061	tacaggcgcc	cgccaccacg	ctcggttatt	ttttgtgtgt	gttagtagag	ctggggcttc
12121	accatgtgtg	ccaggatggt	ctcaactctc	tgaactcctg	atccaccgcg	ctggggcttc
12181	caaatgtgtg	ggattacagg	tgtgagccac	catgctggc	ctatattttc	ttttccaaag
12241	attttttaag	gcattttgtc	ctgtctcttc	ccctgttaag	gatttatagct	tttttaaatg
12301	caataatgac	atactgaac	taagaatttc	cacacagatt	tataagtagat	ctctgtttac
12361	cagatatatt	aatctaacct	gaaaacattt	tttagtactt	tattttgcagg	atgcttttta
12421	gcataaatgt	tgttggaaat	aatttggcca	tttcaacagt	ctttatttaa	atgtttgtta
12481	taactaaaat	aatgtttaat	agacggtgtg	cttttccctc	aaatagaatg	ataggccaat
12541	gcaagaatgt	tagagatcaa	ctgtgttgtg	gatgtttctc	ttttgagaag	atggacatgg
12601	gttgtgagtc	gtcttcattg	aatcccaggc	actagggagg	ctaggcgagg	agaaatcgctt
12661	gaaccaccaa	tggtggagggt	gcagtgaact	gcactccagc	ctgggggaca	gacgaagact
12721	cagtcctcaa	aaaaaaaaaa	gaagaagaag	aaaaaagaaa	agaaattgtaa	cattgtaaaa
12781	tgtatgtccc	tctggaaggc	agtatgacag	ctctcaaaa	acgaataaact	agaatttcac
12841	tattgatcca	gtaattccat	tactgggtgt	ataccaaaaa	gaactgaaag	acagaataca
12901	agcatatatt	tgtactctcta	tgtttatagc	agcattatc	acaaacagcca	aaaagtggaa
12961	gcaacctaaag	tgtccactga	aaggtcaata	aataaaacaa	tgtatgatat	acatacaact
13021	gaactttaat	cagcctctga	aaaggacatt	ctgatcatgt	gtacaacata	acgaactctc
13081	gaagacatta	tgtcaagtga	aataagtcaa	tcacaaattg	acaaattattt	attattccag
13141	ctctacaggg	tctctagaat	aatcaaatat	atagaaaacag	aaattagaat	gaaggtgtta
13201	atgggctaga	gaaattgggaa	aatggggagt	tagtatttaa	tggtgtacaga	gtttcactat
13261	gggaagaaga	aaaaactcgt	agatggatgg	tggtgtatgg	tgacacaaag	gtgtgaacat
13321	cttaagtaca	ctgaactgta	cacctaaaaa	tggttaaat	tctaaattat	aattatgaag
13381	tgagcatct	ctaacatagg	ccatatttcta	gagattaaaa	aaactactgg	acacaggagg
13441	caggcagcgt	ggctcacacc	tgtaatccca	gcactttggg	atgctggagg	ggatggatca
13501	ctgtaggcca	ggagttcag	accagcctgg	ccacatgggt	gaaactctat	ctctactaaa
13561	aaatacaaaa	aatctagctg	ggcatgggtg	cacatgtcta	taactccacc	ctcttaggag
13621	ctgtgctagc	gagaatcact	tgaaccaggg	aggcagaggt	tacagtgaac	cgagattgtg
13681	ccactgcact	ccagcctggg	caacaagagc	gaaactccgt	ctcaaaagaa	aaaaaaataa
13741	aaacttaatta	gaacatactt	aatgacactg	aactgtcac	ttaaaaatgt	ttaaaattct
13801	aaattattaa	tatgaaatga	gcattaaact	agacaggcca	tattctagag	gattaaaaaa
13861	gtcacataaa	atttttaaat	gtataattca	tatgaattat	ttctctgac	tataatgaaa
13921	ttaaatttag	attctagaac	aatatctaga	aaatccctca	aatattttga	aactaaatac
13981	acaagtctga	gtattttatg	catcaaaaga	gaaatcataa	atgaaaataa	gaagacaaat
14041	tcactctgat	gaaaattaaa	acatctcaaa	atctttgaga	agtagctagg	gtagtattta
14101	aatggaaatt	catagcccta	agtggtcaata	ttataaaaga	agataatgg	caaggcaatg
14161	gcctcaggtt	ctacttaaga	aaaagagaaa	caaatataaa	ctgaaataaa	tgaaaggaat
14221	attaaactat	gggttagaaa	tcaataaaat	ataaagcagg	aaaacaaatg	gaagaaaamt
14281	gaaatcaaaa	gctggttctc	tgagacattc	aatgaactgt	taaacctcta	ctgtagactta
14341	tatggaaaaa	aatagaaaaa	acacaaattt	ctaattgaca	gaatgacaga	gggtgtacaa
14401	tattgatagc	tctattgaca	tttaaaaaaa	taagagaatt	ctatttgcaa	atattgttaa
14461	atattgtaca	actcagaaga	aatatataaa	tttcttaaaa	gacacaaact	actaatgaaa
14521	ttcacttaaa	aagaaataat	aacctgaata	acgagatatc	tattagggaa	attgaatttc
14581	aaacctcccc	ttaaaagaaa	ttcaaggccc	acatgtcttc	actagggaat	tgtactaacg
14641	atttagtgga	gaaataaact	taattctgtt	caaatctctc	taagaaaaat	aagagaaaat
14701	tcataccaac	tcattttaca	aggccagcat	atccctgcta	ccaagtgaag	acatttgaag
14761	aaaattacag	accaatattc	ctcaactgtg	aggaaaactt	tcttgtttta	ttagtctctga
14821	gagtttattt	gaagttaagt	acacaaaaaa	ataataataa	cttgtgtgtg	tttcttttta
14881	cttccagcct	gaagtgtctc	aagcaaaaga	aaccaatgaga	tagaataaag	ctggatcacat
14941	agcgacatac	aaactgggat	tctctttatc	atctttccct	ttctttcaat	ctgtttttat
15001	ctctttattt	taactattat	tttaagtcca	gggatcatgt	tgaaggtatg	cgaggtttgt
15061	tacataaggta	agcttgtgtc	atggggattc	atgtgtacaga	ttacttacc	accaggatga
15121	taggcctagt	accatttagt	tatttttccct	gatctctccc	ctctccacc	cttccaaact
15181	ccaataggac	cacgtgtgtg	ttgctcccc	ctatgtgtcc	atagtgtctc	atctcttagc
15241	tccaggttat	aagtgaagac	atgtggcatt	ttgttttctg	ttccgtcggt	agtttgttaa
15301	agataaatgt	tttccagctc	atccatctgc	ctgcaaaagc	atgatctgtt	ttttttttaa
15361	gaactgtatg	tattccatgt	tatatatgta	ccacatcttc	tttatccagt	ccaactatga
15421	ttggcattta	ggttgattcc	ctgtctttac	tactttgaa	agtgctgcga	tgacacataga
15481	ctgtgcatgt	tctggaagtt	gaatcccaac	atatagaaaa	aggtaaacaa	ataaaaggtg
15541	gggtctatcc	caggaaatgca	agtttagttg	aaagtccaag	aatcaattaa	tgtaattctat



15601	tataataata	gactgaagaa	aaaaacataa	tatcttacc	atacagcgag	aaaagtgggt
15661	gacataatca	acatccatcc	ttgatcttaag	ttaaaactga	aaactctcag	caagcttaaa
15721	atagacggga	acttccctaa	ccctaaacaaa	ggcatgttagg	taaaactctac	agctgcagca
15781	cttagtagtg	aaaaattaaa	tgtcttcccc	taagttcagg	aacagaaaaa	gactatccac
15841	ttctcatgatt	tcatttcacg	gtcgtatttag	atgttctagc	caatgcagca	agctgagatc
15901	aaaaataaaa	agacttccat	attagaagaa	aaactatggt	ttttacaaaa	taactatgagc
15961	atctgtgttg	aaaagtcaac	ggaaataata	aaagttactag	aactataaag	agtttttcag
16021	agttctacga	tttgaactta	gtatattaaa	attaatgtta	tttgtatata	catgctaaaa
16081	ttctcagaat	tgaactctaa	aaaaacagta	tttaaacatg	cagcatagaa	tgtataaatc
16141	tttagggtaa	atctcatgaa	aaagtgttaag	atctatacag	aaactattgc	aatcatctta
16201	tttttttaatt	attttatttt	attattatta	tacttttaag	tttttaggta	catctgcaca
16261	atgttcagggt	tggttacata	tgtatacatg	tccatgttg	gtgtgctgca	ccatttaact
16321	cgctcatttg	cattaggtat	atctccctatg	gctatccctc	ccccctcccc	ccccccaaaa
16381	cagtcctcagg	agtgtaattg	tcoccttcog	gtgtccaagt	gttctcattg	ttcattcccc
16441	acctatgagt	gagaaacatg	agtggtttgt	ttttgtctct	tgcactatgtt	tgtctgagat
16501	gatgtgtctc	agtttcatcc	atgtccctac	aaaggacaag	aattcatact	tttttttgcc
16561	tgcatagtat	tcocatgggt	atatgtgcca	cattttctta	atccagctca	tcattgttgg
16621	actgtgtgtg	tggttccaag	tcctttgctgc	aatacattgt	taagagagat	taagtaagac
16681	ctaaatcgatt	ggaaagatc	atcagttcca	tgagttggaa	gaattgtaat	agttccaact
16741	tggtgtacatt	accacacaa	tccaatgcac	ccaggcaaaa	tccaagttag	catgttggtta
16801	gatatactaaa	acctgaattc	aaaattgggt	cagaaatcca	aattgtctaa	tgacgcaaaa
16861	acaactctga	aaaaagaaaa	gaatttgaaa	agactctcag	taagaacaa	atgaacaaac
16921	gtcaaatgtg	aaagaaactg	aaacaaaaga	ctacatgtta	taagatttca	ttttatttaa
16981	agactctttaa	aaacacagaa	ctatagtcat	ataagcagg	tgatrgttg	cacgggtggt
17041	agggcatttg	aggagagaat	taattgcaca	gggacacaaa	tgatagttaa	aaattttaca
17101	tcagatgttg	tgatgtgggt	ttcacccact	gtctataatt	gtcaagttta	ctcaaatgat
17161	atattttaaac	taatgaattt	tattatacag	tattatacta	gagctaaaca	aaataaaagt
17221	gaaaactctc	tggtgggaaa	ctatagagaa	acataactc	ttatgtttg	tgtgttaagt
17281	ttcaggacttg	ttcaaccact	atggagagaa	ggtttagtaat	atctagttaa	atccacactg
17341	catatgttct	ctgactcaga	gtttctcctt	ttaggaaact	attccaatga	taacttgaca
17401	aaaaattgaa	atggtgtaga	cacacagcta	ttcatgtcca	acagctgctt	aaatagatct
17461	tataccataa	attaaagaat	aataaataaa	caagaataaa	tacagccata	acaaaaagac
17521	acaacaaaaa	gtagtccatt	caggtgggtc	atactggcca	aaagtggaa	aatctgagca
17581	tcgattaaaa	taacggctgc	agtgatttga	aacctatcag	ttttgtttaa	cccatgtagc
17641	tcaaaatgat	acttgaaatg	tagctaactg	gtcatctcag	agagatgtca	tgagggcaca
17701	tcattattctc	gaaaactgat	aaataaaagg	aaaagtacag	catattcccc	ccccagtctg
17761	aatatactat	aacttgagggt	aattgaatag	ttgtggaaaa	ttattttaga	aaaaattcac
17821	ctaagaataa	cgaaaaagta	tgatggaatt	ataatgatc	tattttaaac	ccctctcgaa
17881	aaaaataatt	tattcaattt	tcattgtgtg	tgctaaaaa	attagggtga	gccagggaca
17941	gtggcagggt	ccagtcgtcc	cagttactct	ggagactgag	gcaggagatg	cacttgtaac
18001	caggagtcca	aatccaaact	ggacaataata	gtaagacccc	catctctaaa	aagttaaaat
18061	aaaatttata	tggaataatt	tgatgcagat	tatacaaatg	tatagacctt	agtgaaaaaa
18121	ctaaatacca	ctgatcaacc	tgagcattat	taaaaataag	aatctgacat	tgtgggtctc
18181	ttgtttatac	cagtaggag	tatttagcat	ccctacaaaa	atatctctgc	caaatgtcat
18241	acctaaatct	ccccaagcca	ataggtctta	Yaaactttca	caagataata	agagaaaatc
18301	taaatattgt	taaacaaaaa	catagagaag	ttgtcagaaa	agtggaagtg	ggagaaaatc
18361	tatagaaaaa	tggtcaattt	ttacaacaaa	aaaaatgccg	agagagagaa	agagatgacc
18421	atagggtttaa	aagctgaaga	aaactttaat	ttcaagttaa	ggtagtttgg	agccgaccaa
18481	tgctctgtct	gagaaaaatc	aaaacatttg	gatcatctct	ttctctcttt	tttttttttt
18541	tttttaattg	taaaacatca	gagaaactgc	aaagtaaccc	agactggag	aaacaattgt
18601	agagaaagta	gaaaacatgt	ttcagccatt	aatctccctg	ggaaatttaa	tggtttctgc
18661	tatagagaaa	gagccttaga	atctagggtta	cgccacagca	atctaatgtg	gcggagatac
18721	gagaaacagg	cagagcattt	aacagttttg	tcaggctaga	gagacaaaa	tgacagcata
18781	aagggacagg	atccagcgag	gcagttgggtg	aaagatttaa	ccctgaaact	ggcagatttt
18841	tcctttcaag	cattttgtgc	attctaaata	tgatcatggt	aggaatagag	aagttaggca
18901	gaaaactctc	gaaaagcaga	gtaggttttc	caacaatctc	agttaattgag	gaatttaccg
18961	ttaaaacctt	ctagatggag	gggtgcccaa	caaacaccca	tgaacctcag	ttggaaacct
19021	gggaggcctt	tatactacaa	tcaaagggtga	ctgagaggtt	aaataacaaa	atcccaactc
19081	aaacccagta	ttgactcatc	tcacactttg	aatataatgt	atcaggccca	catgtgcagc
19141	caacgcagaga	aaagactaaa	ttctttcttg	aaagaaaaaa	catcatctct	agtcctatac
19201	atctactagg	tcaccaaaag	aaaagagagt	gtaaaataaaa	gtcaagaaaa	atacaatgat
19261	agaaatacac	caacaagtga	tcccaacttt	ggagttatta	gtaaacttat	tttaaatgtc
19321	cttgaatgat	atctttaaga	aaatatatat	aaatttacca	gagcttttgg	aatctatttt
19381	ttaaaaaaca	aaggtaaat	ctaaaactga	aaaaataaaa	taactaaact	taggaactca

19441 atacgtgttc tgagcagatc agacacagaa gatgaaatc tgaactaga agaaagagca  
19501 atagaaacaac ggtgacatga tttttttgtg tctgtgtgtg ttgtgtatag agacaggtgt  
19561 tctgcattgt gccaggctg gtcttgaact tatggactca agcaatcttc ccaatttggc  
19621 tctcccaagt gctgggatga caggcgtgag ccacacaccc tctccagaa aacaggtgac  
19681 tgaagaagag aaacaactaa cagaaaatag agaaaactat aagatggta caggtgagag  
19741 tgaattggca aaataataat ttctttagaa tcacagaagc agagaagcga agaaagcgca  
19801 aaagcaaaac tcgaaaagat attagccaac tctgatacat agaaagggaa ccatactgta  
19861 tttaggcaca tcatagtaaa actgctttaa accaaggaca aaaggggaaa tttaasitt  
19921 acttagagaa agaaagaaat gttaaaccac aagcaataac aataagactg agctattact  
19981 ctgcagacaa atgtgaaat ttaaaacaca atgatacaac tcttaaaaga acattgtgac  
20041 cactctagaa ttctatcccc agtgagaata gcaattactt ttttttttga gacgagat  
20101 ctctttgttca gccaggctag agtgacgtgg cagcatctca tctcactgca acctctctct  
20161 cctgggttca agcaattctc ctgcaacagc ctctgagta gctgggatta cagcatcca  
20221 ctaccatgtc tggctaattt ttgtattttt ctttagcaga gatggggtt cccatgtac  
20281 cttttttttt tttagtaga gacagggttt ccccatgttg gccatctgtg tctcaaacac  
20341 ctgacctcag ttgtctcgcc tgctgggccc tcccaagt tctatggctg ggcactcgca  
20401 cctggcgcaa aataacRttt tttaaactgt aaaaataagg cagcacctt tctgacaa  
20461 actgaggtaa ttgtctcgca gcagacctt actaaataa cgttatgtt caaccagaag  
20521 aaaaatgacc catatggaaa cacacatgtg tggaaaagaa tgaaaaacac agaatatgta  
20581 aatacattag taattgtaaa gtaattatga cagtacaat atgtggcaata atattttgtg  
20641 aaacttaaaa tatttataga attaaaatcc atgtgtgcaag aagtggaagc aggttagtg  
20701 gtaactgggg ttaaaatgtg caaaggtttt aaacttagag aaatggaaaa taatataag  
20761 cattaaacta gggtaagtga aggatgccag ctataatgtc taataaccac taaaagacta  
20821 ttaaaagagaa taattaaatg aatgtagcaa acatatttga ttacttccaa gaacacacag  
20881 acagagaaaa agcaaaacaa agcaggttatg ccaaatagaa aataagacta cagacataaa  
20941 agcaagtata ccaataatga cattaaatgc catgaactaa atgattctat ttcaacacag  
21001 tactttgtat gttgaattga aaacaaaata caactcttat ttgctaaaga gaataatact  
21061 ataaaaaact aaactcgaaa aagttaaaag taaaaagata cagcaagaaa accaggtctc  
21121 actcctacaa cttgcaagga actgaatttt gccacaaccc atgtaaactt caaaaagaa  
21181 tccaagtatc tggatgaaat atagctctgg ccacacctg ccttaccctt ccaacaaagc  
21241 caccatgctt ggactcctaa cctacacacac tgtgagctaa tacatgagtg tcatttttat  
21301 ctgctgtatt ttggcaatt ttgtacacag caatagaaaa ctaatagcta accgatgtg  
21361 ccatccaaaa attaatat atgtttgtat atattcataa tatgaaagt cgcacagcaa  
21421 tgagataaaa tgcattatgt ctatatgcac taccatggaa gagtctcagt acataaagc  
21481 agacaacaaat aaattagaca caaaagagta cctgcatgaa tctatgcatt tctagaaagt  
21541 gcgtttgttg ggtttagta taattctcata tttcttttga taaaactcta tatctttcat  
21601 aagattctca agcaaatcat gaccccaaaa aggcaaggtt cactgatgta gattatgtta  
21661 actatgtact taaaaaaaat taaatcagca taattgcctc aattttttat gttgatgtg  
21721 catgaggaca gaatgcaatt tgtattgttt ggtatctttt aagtttttct tctgctaatt  
21781 atcatcctta aattcataaa aaagcaacaa agaaacttaa tgcattggaa ttcaaacactg  
21841 agYgcaagtg taaattcttt gcagggttta tcaattttgt cttcaagaaa ggaatcaactg  
21901 cctgagcatt tggctcagg cctgtaatcc cagcactttg agaggtcaa gttgggaagaa  
21961 cacttgagat caggaaatca ggcccagccg gggaacata gcaagactac atcttataat  
22021 aaacaaacaa attagcagg cagagtatg tgcactgca gtcacagYt ctagggaagc  
22081 tgaggcagca tgcactcag gactcagaaa ttcaagctgt cagtcagcta tgcacaccc  
22141 actgYacccc agcgtgaaca acagagtaag atcctgtgtc ttgaaaagaa aagggaaagc  
22201 aagaagaaa tccaatgttc tgcaattctt tcttattccc tccagctcca atcatgaatg  
22261 tgcgtctcgg ggttggctag attcgaggac accaactagg ccttctcact gatgggtaga  
22321 accaatgcaa gtggtaaagc cttgagcttt tctctcact ccttttctgc tctctttcc  
22381 catgtgacag caaccaagga ttgggtcatt tcaggagaat tattttctgc agaggttcc  
22441 atcagacatc gcgaacaaat atattctcat tatgtactat ctttctctct ggtatctct  
22501 ctctgatatt gacagatgYc accaagtatg agctagctYc aaaaaggatg tcttataaa  
22561 cactgcctgg gttactaaca tatactgcac caaaattgtt gccttcaatc cactaggttaa  
22621 taaagtgtta ccttggtttt cttcattcat tgtttctgac acaggtattt ggtatttctg  
22681 ccttgcaga gaacatggg atggtcaggt acaggtattt ggtatttctg gatccaggt  
22741 tataaaactg tcttcatact gtaagaaaat ttctgctaat ttctgctgag ctataggta  
22801 taccacactc ttgcgtatta agacatacag tcttgactta aaacatattt attagggat  
22861 tggatttctca ctcaaatagt ttaggaacat tctcagaa atgtgcattt aatgagaata  
22921 tcatatttat gttacactct cctgaaagaa atgtcatata ttcacatagt gctaacatga  
22981 caccaggaact atgtttgaat gtgtttgtag aaaaatgact tgacatcaga acttgctctt  
23041 taagaccagg aatttcatat tcatacttag aacatcatag atggtatata ggcagaggtg  
23101 atagaaaatg ctaaaaagtg ctttacagag agcagagagt gtgtttgaga atgtgagct  
23161 atttttttaa atgtaaaaac caaatgacca ttgctcttag tccaacttta gctctgggac  
23221 aattatcaag cctttKact ttctgagtt cctttKact tgcagggtt cacttcca

23281 ttactgcccc aaacacatca gogaactggt taattcaatg atataagttg tattttttgta  
 23341 ttaaaaataa cacccttttg attgatgact aattcaggta agaaaagctt ttgtgttttg  
 23401 gRtaaaactc atcaattacc taatgacact gaattcaggt cattagcagg gaaggaagaat  
 23461 atttctaggt gttagagaaa ttcattttta ataatctakG ctgttttttt aggtataacta  
 23521 cctttctctat tctgtaatfac acaacaaata ttggatttttt cctgtttacac accattcaac  
 23581 acaaaacac acacaaagc aatcactctc tcaacaaacR tgaactctgg tgtctctttg  
 23641 attcattcca atttctgact atctacctgg agatagcagt ttgagggtctc agtcccaag  
 23701 gactgcggt ttacatgcca aatgcaagta ataggttatc actcttagtt ctcaacctc  
 23761 agctgtgatt cagagttttt atgacctctc cctcatggtt gattaaattt cttagagtga  
 23821 tcacagaagt cagagaaata tttatgttta cccgtttatt ataaagaaca Wtggaaaggt  
 23881 tacagataaa agtccagatt gaagagatgc atagggtgag gtaaggggaga agggatcacg  
 23941 agcttcaaaag coactctcag gaacctccac atagtcagct ctccagaagc tccccaacc  
 24001 aagatatttt tagtttttag gaagggtcca ttacacaagc gtgatggata gaattattgg  
 24061 ccactggcga tcaattcaac ctttagctct tctccctcc ccagaggttg agatgtgggg  
 24121 ctcaaaagct caacttctca atcatgccc gtgtttttct gtaaccagcc caactctga  
 24181 agctacatgc gggttcccaa acctagtcat ctcaatagca tgaactaaag cttcttaga  
 24241 cactctgcca gggtccaggca cagtggctca catctgtaat cccagcactt tgggaggcgc  
 24301 agggcgagat tcaacttgag gtcaggagtt tgggactagc ctccagcaac tgggtgaacc  
 24361 ccactctcac taaaactgca aaaattagcc aggcgtggta gtgcaacctt gtaattccag  
 24421 ctactcagca ggagggagca ggagaatggc ttgaaccag ttgacggag ttgcagtta  
 24481 ccaagatcat gccactcac tccagctggg gtgcagagaa cctctcgcta aaaaaaaaa  
 24541 agaagcagta tgcacggaaa tcaaacacca aatatatact tcaaatatc gacctcaac  
 24601 gatggtatag attatgtgtt accattcaaa aaataaaaaat aagctctcaa ttgaaaaacca  
 24661 gctctagata caattttgaa aaagcaata tctcccact attttttagt agtggagtgt  
 24721 atttagaattg ctgtttcccc tctgatata attaaatggc tcccatttca aatccatttt  
 24781 tttgattgag gatactggtt aatgtttttt ctgactctaa atatagtttt ttacatgga  
 24841 aggccttgaa gtaactctct cagtcaaaa gatactgga aatttttaggt tgaactcaagc  
 24901 tcaatttgaa atgatgttta ccaagttaat attataagca aaagttagta cctctgtata  
 24961 aagagaattg tgcaaaagctt aacaagattt taaagtatag gttgagtata cttttatgta  
 25021 aattttttga gccaaaagtt ttcagagttc agatttgggt aattttttga atttttgat  
 25081 atactctaag agatactgtt aagatgggac ccaagtataa acacaaaatt cattaatttt  
 25141 catacatacc ttgtacacat aggcctcgag taattttata aaactcttla aatattgata  
 25201 ttaaacatga aacagagttt gtgttaagca ctttggatgt gaattctcca cctctggtat  
 25261 cagtctcagg ctcaaaaaaa ttttagattt tttgattttg gatttatgatt ggaatgcac  
 25321 cctgtatata tgtatataca ggtatataa tatataata tatataata tatataata  
 25381 caccctcata aaaaattttt tcacaaatgt aatataaaatc aagctcYctc aggtttttct  
 25441 ataaaacttt tttaaacgta gctatcaaat gcttctctgg attaagtaac attatctgaa  
 25501 atttttttta cctctactca tttttgtctg taaattatga ttttaagtaac aatgatgttg  
 25561 gccaaaatta agggagcttta taagaaactg gcatatgaaa tttattttat ttattcttga  
 25621 cactactgtt agggtaattc tgaactccatt aactccatta aatagctctt ttgtttttg  
 25681 ttgtttttgt ttttgagacc gggtgtgtct tttgtccagc gctggagctc agtgaacaaa  
 25741 tcatagctca ttttaacctt ggagctctgg gctcaagcta tctctctgct cagcctctg  
 25801 ggactctacg ccataatgcca ctgacttgga cttttttttt tttttaattt ttgtagacat  
 25861 gggtcMtgcc tatgttaccg aggcgtgtct tggactctca gccctacaca atctctctg  
 25921 ttttaacctc caaagtgtct ggattacaga ttgaaacca cacactctgc ttatctctgt  
 25981 atttgttcca tccgccaata ctcaaccaat aattgatata aaaaaaacat ctatagctac  
 26041 caacatcata cagttttttt atttagcttg ctatggcaac ggagaacata cccagggaac  
 26101 cattaggagc attctagtaa agtgaacata gaaggggctt ctatggact ctgacaagt  
 26161 gtggggggc tctgagaaca gattatgaaa gtaagtggg tgcattcaga gaagagaggc  
 26221 aattcgataa ttggtatctc aatgatgtcc ccagagaaca ggagacatgt caatcaaat  
 26281 cacattgtca ttgcaaggaa gcagagatca ctcaagaaag gggagtgtgag tcattttatg  
 26341 gtgcagcag agccttgctt ttgttttagct agaaaactg ttgttactgg aagcatcact  
 26401 aaatttctct tgaccatgtc tgcattccat gaagaaaggt gttgtttttg gttagaacca  
 26461 tcaactagtg caagcagggc acatttttgt ttctcacga atccacagta ttgttatgag  
 26521 attagtttct actaccagag Rtgcgtgtgc ttctctcaca ccaactgttt gtttMatct  
 26581 tattaaggct agaaatttga aaaaagagaa aaagactgac cactctctc aagtgaact  
 26641 tagaacctcc caaggcaaca cagctacagt gcatctattt gacaccaaYg gaatcagtag  
 26701 catcatctca atttctctatM cttcttttcc ctttacccaa tttttttgtc atttccat  
 26761 ttttacttta catgtttttg aaatcacacc aatttttttt taggtctaaa gactcagaa  
 26821 ataaaaaata acrgtcaagc ttggattgtc tttatttgt gactcagaa gaggatgag  
 26881 aagaagtgag gaagtgttta tgaatgtgga tggaaagtac ggacaatgtc gggaaacaa  
 26941 ccaagtataa ggaactgtacc ctttctcaaa ctgttatact ccttgtgac  
 27001 ctacagcaag cattgaactt ctttaagctt tattttctct atgtttaaa gataataata  
 27061 ttgtaccaat atagttagga taaaaaaga attagcatgt gaagacttc tRtaatttt

27121	acatgacaat	gtatgaatat	tttgttaaaa	atttagtgcc	aaaYtctttg	tgcacaaaag
27181	gcataYtglt	taaacaaaact	ggatctacc	tatcttagaa	atttttatta	ttacacagatt
27241	taaaagccaaa	tatacaaaa	tacgcaaaaag	taggagaagt	ggtttctaaa	tttttggaaa
27301	acaaataagat	gagtctcttaa	gatgcataga	agggcaactga	atltcatgac	agaaatagta
27361	agaagaagaag	aaatagaaaa	tggtctcttgg	ccggcgccag	tggatactgc	ctgtaatccc
27421	agctcttttgg	gattcaacaag	cagggtgggtc	acctgtcaga	agltcaaacg	cagcctagcc
27481	aaatatgata	aaacctctgt	ctactaaaaa	tacaaaaaat	agctggggcat	gagggccagt
27541	gcctgtaatc	ccagctacttt	gggaggttga	ggcaggagaa	tcgcttgaa	cttgagaggtg
27601	gaggtgtgac	tgagctcgaga	tcggcgccact	gcactccagc	ctgagccact	gagggaaact
27661	ctgtctcaaa	aaaaaaaaga	aaaaaaaata	aaagaataat	ttctcttact	tgaaaggtgac
27721	tatatgaaa	tcttttacct	ttctttctac	gtagaagaac	agaatacaca	ttatacaga
27781	attatgcaca	acctcccatg	ctcataaata	ctaactcaat	gagccctgat	tcagattcta
27841	aatcatcctaa	ttaaatgtcta	attggcttat	tctctttaaa	atgagtttac	actgattcta
27901	tctatttccct	gattctttgga	cgatgagtc	tagtttccct	gattataaaag	gacatggct
27961	ggatgtgtact	ataagaagaag	caagcacata	aaatcatcaga	gagccccatg	gccaaatcac
28021	catataaagta	tctatgaaga	cgcaaaaata	acttttacc	tatatgctgt	gattttatag
28081	taacataagt	attgtatctct	acaaagagtc	atagagacac	attgtgtctca	tttcttact
28141	gtcataactgt	tcaaaaattta	tatgcagttc	attcttgcaa	attgtctcca	gagacggtt
28201	agaggcagaca	tacgataacc	atttttgcac	cagactctca	tctcataatc	aaacacacat
28261	agcctctcttt	tccacagtc	ccctctcatt	aattaaaatg	ctcttgtaag	aaatgacctt
28321	ggtgtgtttac	aaataccaaag	Wctgcctcta	agagcttaag	ctttgtctcc	attgaagaca
28381	tgcagacact	agctccagaa	tggtctatgg	ttgagttctg	ctaagtctgc	acctgtatgt
28441	agtatctctca	gagaagaagct	ggctgttgca	caggaggcaa	tgctcccttg	aaaaggaccc
28501	agctctctat	agctgggaaa	agtcacattg	cttcagggtc	attttttgaa	gaattgttgt
28561	atagcctatg	cataaaacttc	tgttaaaaaa	tttgcaaaa	caactgtatc	tttctattct
28621	tagctcaaga	gttcaactca	ttttatttaa	aaatatttaa	aatatgagtg	tgcagaaagt
28681	ctgtgtattac	aaaaacctga	tattcaaacg	aaatgagtg	aattaaagtg	tgaagagttg
28741	gcttccattc	acattctcca	tgaagccaa	agacacaa	tctactcaga	agatttctga
28801	tcttccccag	taatacattt	aaaaggtgac	tcagaccagt	gtgtaggtct	atgtttatct
28861	ttctgtccat	ctcaaacacc	ccctgtgagc	aggatcacaga	taaggaanaag	gagagattgt
28921	ttcactagtg	tggcaggcaa	agctcagata	actgattgaa	agatgccatt	attggacata
28981	ggctgtgttct	agcctacctca	gttgtaaggc	acacgataag	ataaaaatcc	caggaaagtg
29041	acatttttga	aatcctgtct	tgactcagt	ttttcaacag	gcattgtcat	gggtatgaga
29101	agagaaaaaa	tggggagaaa	acgctagaag	aaattaaaat	gatgtgacat	gatatctacR
29161	taaggttttc	cttgacatag	acctgtgcag	gggctcttgc	agatatttat	gacactgtgg
29221	atgccatgtg	tctagtggaa	cttccagagt	tctgtgaagc	cacaccagca	tctgactctg
29281	gagtttctat	agggccatgg	ctagttaggt	gatagattct	gggtgactgt	aaactgttat
29341	acagccaaagg	ccatttttgg	tgtacatgta	gggtgtgtgc	tatgtgggtg	taattctcga
29401	tttttactcc	cacaatatgc	tgagtgatca	tgaagaatac	atttttaaat	tgaagccaga
29461	gtatcaactg	gttttatcaa	aacctaatg	gaanaagtat	atatatatat	atatatatat
29521	atatatatat	atatatatat	atttatatat	attccacctc	tctaggtatat	ttttgtgtgt
29581	ttttataatt	gttattttca	ttgtagttaag	agtattttaa	tgaagtctac	cttttttaaca
29641	aattttaagt	gtacaataga	gtatcattga	tgaataataa	tacattattta	aaaagatata
29701	tggtgtgtgag	gttatgtaga	aactggaacc	cttgtacatt	gtgtgtggga	aatgaacata
29761	gtgctgccac	tacggaaaaa	agtatagagt	ttctcaaaa	acttaaaaaa	agaaactgtg
29821	tatgatgcaga	caatcccatg	gatYggtatt	tatctaaaag	aattgaaact	gatctttcga
29881	agagatagta	tttgaccccc	tattttcatt	gcaacattat	ttacagcagc	cagagttgtg
29941	aaaaagtgtc	attgatggat	aaatgtgtaa	agaaaatgtg	gtatacacag	acaaggaatt
30001	ttattcagcc	tttaaaaaaa	ggaaattctg	caatgcaaca	acatgagaag	accgggagga
30061	catctatact	agtttaataa	atgagacaca	gaagggccag	ttctaaagct	tcccgggcct
30121	gaagagctta	gtagatgaaa	aactcctccc	ttctcagccc	cagccggaag	acgcaaggcc
30181	acttgcgtca	cgacgctgag	tcagcaaggt	agcagaagac	ggaagaagac	cagcaagaag
30241	acacctaacc	tggcaggaag	acacgtaccc	gtgaagatag	agaaagagcc	acatccagta
30301	caacgtagca	gttaactcag	aataggacac	ctcctgtttc	caggagacac	taaaaactct
30361	gccccctccc	cacttggggg	ctgacgccat	taggcctcag	cccgctctca	ccccagcgct
30421	catataaaca	gcattgtgac	cgcgccgggt	gtttactgct	gtaatccagc	cacttctgca
30481	ggcctagggc	ggcggtatct	gaggtcagga	gatctagacc	atccttgcta	cacgggggaa
30541	acctgtcttc	tactaacaat	acaaaaaatt	agccgggcat	gatggccgga	gcctgtatgc
30601	ccagctactg	gggaagctga	ggcaggagaa	tgggcgcaac	ccgggagagc	gagcttctga
30661	tggagKggaga	tggcgctact	gcactccagc	ctggggcgga	gagcagagc	gacttctcaa
30721	acaaaaaaaa	aacagcatgt	tgctccacac	tgcccttgtgt	gtctgtgtgt	cgactcttga
30781	gtgttcgaac	gtatacaaga	accttaccac	aaatctgtat	gattcttttr	gattctgaat
30841	gtaaaacagt	aaaaactctta	aatgcagaga	gtagaatgtg	agttactagc	gactaaaaag
30901	aggggggtaac	tgggaattac	tattcagcag	gtacaaaatt	tcagtttaac	aaaaatgaatt

30961	agttcttagag	atctgctgca	caacactatg	cctattgcca	ttgtcgtttt	cttgttttgg
31021	tttgggttgg	ttttatgctt	aatttcttct	taaaaaaat	kattgactca	agcccttaat
31081	agctgtctgg	gaaatctatg	ggggaacaaa	gaactataac	acaccagctca	cccgctctgt
31141	ccctttccca	ctctgtttga	gtgagatgtc	ttagatagat	ttgatctcat	cgctccaacc
31201	tgctcaactg	ttatgcttag	catttccaaga	tccttagcaa	gacataaacg	ctctgtcaat
31261	gtggctgcta	ccaaccctc	tagctgcaac	tcctgcctct	tcctataatg	cagcttagct
31321	tccaacagta	ctgaaccact	gaatttctgt	aatttgataa	cgattttaac	cttctctggc
31381	tcactgtgca	ttctttggtc	agaatccatt	tcctggctct	cttttttctg	tgtataaatg
31441	tagctgtatt	ccatggctgc	ttattttacat	atatcttctg	ggaaagctgt	ctttttaaaa
31501	tttttgcaaa	tttatgggtt	atRgtgact	tttttatgt	ttttatagca	tggaagttag
31561	tcaagctcag	gtattcagg	tgctccatc	cccgatcaaa	tacgtttttg	ttactataag
31621	tcactctgtc	atcaaacatt	ggattttatc	cttctattta	actgtatgtt	tgtaccattt
31681	aaccagctct	tttcaatttt	cccgctctcc	ctgactcaac	tttccattcc	ctgtttgtct
31741	gtcttttact	tctactttta	ggtagccaaa	ctttttaagt	cccacataac	actgaagaaa
31801	tacaataatt	gcttttttgt	atctggctta	tttcatttaa	gataaagtact	ccagttttgat
31861	gtttatgcaa	atgacattac	ttcattctat	tttatggctg	aatagttttc	ctttgttttt
31921	atgtatgcca	cattttttgt	agccattcat	catgtatgga	catttaggtt	gatttttaat
31981	ctttgtctat	ctgagttagt	ctgcaataaa	catgttagtg	caggtagatg	tttgataaac
32041	agattttctt	ctgtgaagcc	ttttgtgatt	taacctccag	tttcatctat	tagacattct
32101	ttatttactt	accgattatc	atttcaaaaa	catatcatg	gactccctta	ataaaatgtc
32161	tcctactaca	gggttccaag	ttgctgggga	ttgaaggggt	ggaggaatgtg	ggacttccaa
32221	tgcataaaat	gggacagtga	aaaccaggat	agttggctac	cttaccatgtg	atgggacctg
32281	accttctcgt	tagtcaactt	tgtttaaaat	atatttcccc	cattatRgctt	tgatgttctt
32341	tgaagRcaat	gtctctcYgt	tggtccttgg	catataaact	acttttgtat	gctttaaatg
32401	tattcagtga	atgattttct	attttcaaa	ggcttaaaat	agctctctgg	gtatgaacat
32461	atttttttga	tggtttgttt	ccttctgttt	aatagtagca	ctgtggaatg	ttttggctgt
32521	tgccccacc	aaatcttatt	ttgaattccc	atgtcttctc	cttctgtctt	tcagtataga
32581	gagttaagct	ccaagaatct	gatggttatt	ataaggggga	actttctctc	ccaagctctc
32641	ttttttgtct	ctgccatcca	tgtaagacat	gactggctgt	tccttgcctt	ccacctgtat
32701	gttgtagcgt	caccagccgt	gtggaactgt	aagtccaata	gactctcttc	ttttgtaaat
32761	tgccccgtct	caggtgtgtc	tttatcagca	gcagtgaata	ggactattac	agtaaaatgg
32821	tacaagtaga	gtgggtgtct	gctgaaaaga	tacgtgaaaa	tggtgaagca	actttggaac
32881	tggtgttatg	gcagaagtgt	gaacagtttg	gagggctcag	aagaagcacg	gaaaatgtgg
32941	gaaagtttgg	aaccgttagag	acttgttgaa	tggtcttgat	caaaatcctg	atagccatat
33001	ggacaataag	gtccagcgct	aggtgggtct	agatggagat	gaggaacttc	ttgggaaactg
33061	gagcgaaggt	gactcttgtct	atgttttagc	aaacacactg	gtggcatttt	tcctctgctt
33121	tagagatttt	tggaactttg	aacatgaaa	agataaatta	gcgtatcttg	tgaaagaaat
33181	ttctaaacag	caaaagcattc	aagagccgac	ttgggtcctg	ttaaagggcat	tcaattttat
33241	aggggaagcag	agcataaaaa	ttcagaaaa	ttgaagcctg	acaatgtgat	agaaaaacaa
33301	agccccatttt	ctgaggagaa	attgaaagctg	gctgcagata	tttgataaag	taatgaggag
33361	ctctaattgtt	atccccaaaga	cagtaaaagaa	aatgtctcca	gggcatgttca	gaggtctctc
33421	tgctcagccc	tcctatcaca	gactaggagg	cttaggagaa	aatggtttca	tggttccatg
33481	ccagggtttcc	tgctgtgtgt	gcaatctatg	aagttgtgtg	cttgcatctc	ggcacttcca
33541	gccatgatac	acaggggcca	aggtacagct	caagctgttg	cttcagagaga	tgcccaagctg
33601	aagctttggc	agcttacatg	tggtgttgag	cctgtgggtg	cccaagaaat	aagaatttag
33661	gtttgggaac	ctctgtctag	atttcagaag	agttatagaa	atgcttggaat	gcccgacaaa
33721	aagtttggct	cagggatggg	gcccctcatg	agaacctctg	ctagggcagt	gtggaaaggga
33781	aatgtgggggt	tgagtcacca	acacagagtc	gctactgggg	gctactgggg	caagacttga
33841	gaagagggcc	accatctccc	agacctcaga	atggttagatc	caccgacagc	ttgcaactgt
33901	cacctggaaa	agctacagac	actcaacacc	agccctgga	agcagctggg	atggagcgtg
33961	tacctgcaa	agccacaggg	gtggagctgc	ccaagaccat	ggcagcacac	ttctgtcatc
34021	agctgtacct	cggtgtgaca	catggagatca	aaggagatca	taagtgttga	taagatttga
34081	ctgccccggc	ccggcgccgg	tggtctcagc	ctgtaatccc	agcaacttgg	gaagccaggg
34141	gtttgggtac	atgaggtcag	gagatcgaga	ccatctgggc	taacaaggtg	aaaccccgct
34201	tctactaaaa	atacaaaaaa	ttagccggcg	gcggttggcg	ggcctctgat	tcocagctac
34261	tcgggagcgt	gaggcaggag	aatggcgtga	accggggagg	cgagactctg	agtgagctga
34321	gattgcgcca	ctgcagctcc	cagtcggccc	tgggcgacag	agcgagattc	cgcttcaaaa
34381	aaaaaaaaaa	aaaaagattg	actgcccacc	aggaatttca	aatatcccaa	aggtcgaaat
34441	ctcttttgtt	tgggcaattg	ctccccattt	ggatggctgt	atatacccaa	tggtctgacc
34501	cccatgttat	ctaggaggtg	actaaactgc	ttttgtattt	ataggtctat	aggtcgaaat
34561	gatttgcctt	gtcttggatg	agatttttga	ctgtggactt	ttgatgtta	gctgaatgat
34621	gttaagacct	tggggaactg	ttgagaagcg	atgatgtgtt	ttgaaatgtg	agcagaaatg
34681	atttagaggg	gaccagcgct	gaatgatagt	gtttgtctgt	gtctccactc	taactctcat
34741	ttgaattccc	atgttttgtg	ggagggactt	ggaggagggt	aatttaacta	tggtggcgaga

34801	tctttcccat	gctgttctgt	tgtatgaaag	taagtctcat	aggatccgat	ggttactgta
34861	agggggagtt	ttctctgcaca	agctctcttt	ttgctgctgt	ctatccacRt	aagacatgac
34921	tggtctgtcc	ttgctctctgt	ccatgatgtgt	gaagcctctt	ctgccaatgt	gaaCYgtgaa
34981	tcacaaatac	ttcttttcttt	tgtaaaatttc	cgactcttgg	gtatgtcttt	atccagcaga
35041	tgaaactgtg	ctaatcacaca	ctgttRataa	gagitttaatt	aaaaaccccc	agaaaMagtc
35161	taaatgtcca	tttaaccaaat	aaatgataat	tacatgttga	ataaatatc	tatggtccag
35221	tcacttggaa	tgatcacaga	aacaaaattt	agttgactac	caatgtacaa	acagatgtga
35281	tgaattctga	aattataatt	tgtatgtata	ttggaagata	caaaagaata	cataataaaa
35341	tactatctga	ataatattta	attatctgcta	aaagggaaaa	gaaaaaagac	ttccattatc
35401	aatcaatatt	tggtgggttt	tattattttt	atcaRtacta	ttaacttttag	aaagcagctac
35461	tactatcgaa	catctcaatc	atgccacaca	tttagccctt	aataatacat	atctltttttg
35521	attttacttt	ttcaataaact	ctgcaactcta	taggtgaaag	aaacaagata	ctgaggtagt
35581	actgaggtta	taaatgcatat	gaggtagtaa	atgccttacc	ttgctttgcc	aaaccaagat
35641	gtctcaactg	gattccaatac	ataaatatgc	tgttatgtgt	aaagccagat	ttcatgtctg
35701	gatctcaatt	atctgttaat	gtgaatgtct	tttgggttaa	acacactgca	ttccattaat
35761	tctctcatct	tccatgatcc	ttgctctcca	gtatttatgt	ctttgtataa	tctttctctg
35821	ttgagttgag	cgaggacctg	ttaactgtct	ctaaccagta	taatttgcaa	agagttgag
35881	atgtcacttc	tgtgattaca	ttacattata	taggattcca	tcttgcatac	agaattattc
35941	tggaagctgt	ctccttaggc	ccaatgaca	gccctagct	caagatagtc	tcagcagtaat
36001	ttccagcaaga	aaagccagggc	tctgagtcct	agaggcacat	ggaaatagac	ctcccaacca
36061	ctcagcttga	aagaggatgc	ttctctagtt	gagctccac	atagaaaccc	aggtccagctg
36121	acggctgtgag	tacagcttta	taagacccta	agcagaggac	ccccgcctaa	tttgtgcctt
36181	aattctctgc	ctaaaaacaca	attaggttaa	aaatgggtgc	tgtttccagt	tgctatgctt
36241	gtgatagtgt	gttatgtctac	aaaagaaaa	gaatacacta	tcttttttaa	aaaacttaatt
36301	tttYcttcaa	agacaagtta	tgcaattatt	atgggagatt	attctacagt	tttataagaa
36361	gccttaacca	gagtttggaa	ctggactgcc	tgccctgatga	tgttaattata	ggaatttaatt
36421	ccccaaaatt	agtaaatata	tgaaaactgt	aagagtacaa	aataatattc	tttgcatac
36481	tgcccaactt	gaagctcagt	tagtgctact	ttgactaggt	atggtgtcat	gagtgaggtc
36541	aggttctcagc	ccagactgag	ggttgagggg	agtggttgga	catggggcag	ggagctggag
36601	gaacactgtg	gagacagcag	gtaaaatgaga	catggcttta	ttcagaagcc	tttcttcagg
36661	gtcagttgtta	catttataca	ctacacaaa	aatagtggtc	gagagccaga	ttggctgttc
36721	ttctatgtta	tgtcaacatg	gctatgatta	tataagcat	gggactatgy	gcctgtgcta
36781	caaatcgtct	gagtcataca	ggctgtttac	cttggcctat	gctcgtctgc	ctatgctgct
36841	ttgctgcag	cacagccaat	ttccttacct	atggcattgt	tgaatttgaa	atgtctatgc
36901	cacactgcag	agagagagaca	tccaccagtc	ctcatagac	acacatagct	tattaaatcc
36961	tgctctcttc	agaagccaag	acctatgcat	tcctctatga	atagctttta	aaggttaatt
37021	aaaagcagtg	agtaataaaa	atttataagc	ttctgcttat	aatttttaa	ggatcatttt
37081	ctgacctttt	gattgtgatc	ccatgatgta	ctatccagat	gccagttgtt	cactgttccc
37141	aacaacttaa	aggagaactc	aacagaaact	agagtctctg	cttttccag	aaagtttctt
37201	caaggttttgt	tctgtgtgat	acaaattctc	atggctgtta	gggtttctgt	ttggaattgg
37261	attattctttt	cttggagatt	cacaatctac	attagcattt	ataataatc	aaaaagtttc
37321	aactaagaaa	aataatatac	aRttctctta	ctagcattac	atgtacttat	ttgtctccag
37381	aaagctttct	caaaaacacc	ttaaaaactc	aRtaactcc	gcttgagaaa	gagttgtgtg
37441	aaagatgtga	ctcatcaatt	tggaaatctt	atttttaaaa	ataattacag	tacgtctcac
37501	acttgaattt	tgaagcatca	aatctattaa	atacaaaaac	aatatcacag	gaagtaaaac
37561	taaaagtcca	catggaaaaa	tttaacaata	aaacattatt	agaacttaag	tagttaattg
37621	tgaccaagac	tggtcaagag	ataaacttgc	acatattaat	gctgttatgc	tcagctttta
37681	tgctccagat	aagactataa	aagaatgatg	tcaatgttag	gaaggaaagt	taactacata
37741	gcaatatatt	tgtattgtca	atttgaatat	tttactgaaa	tatcagcaaa	tgttatatat
37801	tgaatttttc	taacacagaa	gtaaaattgtc	ttctgtcttc	aaatttttat	tgtattcccc
37861	aacaccaata	aggcaggggt	atgYtccac	atgccaaact	agctacagtg	attctttctg
37921	agtttccata	acacagacac	atctcccaac	ttagcaccta	ctgtctttct	gcctgaaatt
37981	ctcttctctc	aaatgtctga	atggcttact	ctctgttttt	atccagttgt	ttctctaaat
38041	atctcatcag	ctcttctctg	gctacctctc	aaaaagKggc	acataccagc	ctcactttcc
38101	tcactctttt	actctttata	tttttcttaa	ttgaaataat	ctatacttga	gaagttttct
38161	gttggtctat	ttgtcttatg	actatttacc	ctagataccc	aaactgcaata	tatgattttc
38221	cacagcagag	attttgttca	gttgtggatc	ccagcatgtg	aaagcagtag	ataattctca
38281	ggaagtatga	aataaataagY	Rcatgaatta	atggaatagc	agaagtctat	ccttcagaa
38341	agacttatgg	cgactcgttt	aaggttaagt	taacatagga	aaagcagtag	aattcacca
38401	attgaagata	aaaaagaaat	atgcaatcat	ctcaatgcat	ataaaaacac	atttatacaa
38461	atctgacata	tttttttaat	gaaaaactgtg	cacacactag	aaatagaagc	gaaccccttc
38521	aaactgataa	tcataaactac	aagaaaagct	acagctaaca	tcacaaatgtt	aaagactag
38581	atgctagatg	ctttctctta	atatggaaa	caaggcaagg	atctccactg	ccagcactta

38641	tattcaacat	tacactagag	gtcctagtca	gtaaaaataag	gtgagaaaaa	aattaataaaa
38701	agccctccag	aaaggaaaaa	gagcaaatga	aaacaagtcta	tttccagatgt	acataaattgt
38761	ctatatagag	cattatatgg	agctctattaa	aaatgttacta	gtacccaaat	gtgaattttag
38821	caagatggca	gagtaagaaa	tgaacaatca	aaaaattttat	gtgtttcttat	aaacatgtaga
38881	tgaataattg	gaaaatgaaa	ttttaaaaaa	ttatctattgc	taccatgagc	caataaaattg
38941	aaaaatttaa	gataaattgga	taaatttcta	gatactatca	agattgaacc	aggaagaatat
39001	tcaaaacctg	acacagaccaa	aaacaagcaa	tgagtcca	gccataatga	aaagtcttact
39061	agtagagcgt	gggtgcggtg	gctgatgcct	gtaattccca	gcttttgga	gggtgaggtg
39121	ggcagatcac	aaagtcaggga	gatcgagacc	atccttgcta	acacagtgaa	acccgtgtctc
39181	tactaaaant	acaaaaaaaa	aaaaaatta	ggatgtgtg	cgaggtgcga	tgtcccgag
39241	tactcgggag	gctgagtcag	gagaattgca	tgaaccagg	agacggagct	tgacgtgagc
39301	tgagatctgt	ccagtcgact	ccagcctgag	tgacagagtg	agactccatc	tcaaaaaata
39361	aaaaataaaa	aaagtctact	agtagagaaa	agcccaggac	ctgatgactt	cactgtcgaa
39421	ttccaccaa	tatttaaaaga	agaactaata	caaatcttat	tccaactatc	atgaaaaata
39481	gagggaaggg	gaatacttcc	aaactcattt	tactaagaca	gtattacctt	aatatacaaa
39541	ccagaaaaag	acaaatcaaa	accagaaaaa	aaatcccaa	aaatccataa	atttatatgg
39601	aaccacaaaa	gaccgacaa	agccaaagct	atcttaagta	aaaaataaga	acacaggagga
39661	agcatatata	ctgccttcaa	attatactac	agagctgtag	taaccaaaac	aaagatgtac
39721	tgccaaaaaa	agacatagac	taattgaccca	gaattagagag	ccctacaaata	aatctacaca
39781	tctatgtgtg	acttgtttttt	gacaaaagtg	ccaagaacat	acactgggga	aaagagctctt
39841	ttcaatattct	gggtgctggga	aaactggata	ttcataagca	gaagaataaa	actagatgctt
39901	tatctctcac	catatacaaaa	aaaaataaat	ccaaatgaat	taagaactta	aatctaagtct
39961	atcagactcga	aaactactaca	agaaaaacatt	gggcaaaact	cccaggcatt	gggtctgggca
40021	aagattttctt	gaaaaatacc	ccacaagtac	agggcaacca	agaaaacatc	gacaaatggg
40081	atcacatcaa	gttaaaaagg	ttctgtgacag	caaaggatac	aatcaacaaa	gctggagagac
40141	aaaccacaga	atgggaagtta	atatgttgcaa	actaccctct	tgacaaggga	ctaagtgtcca
40201	aaatcacatg	ggaactcaga	caactccaca	ggaaaaaaat	ctaataatcc	aatcaaaaaa
40261	ggcgaaaaaa	ttttaatagt	tattttctcg	aggaaaaacat	acaaatctca	acacagactt
40321	caaaaggtgt	tcaacatcat	tgatcatcag	agaaaatgcaa	atcaaaaact	caatgagata
40381	ctactctcact	ccagttaaaa	tgttatagac	agggcaataac	aaatgcctga	aagattttgt
40441	ctctgtacag	tggtggcggg	aattgtaatt	agtacaacta	agatgaagaa	gagttttgag
40501	gttctctaaa	aaactgaaaa	tggaactacc	atatgatcca	ccaatccaac	ttctgggtgt
40561	acacctaaaa	gaagcaaaaa	ttgtatatgt	aagatatatc	tgcaactctta	ttttttgttc
40621	agccctattgt	attatagcca	agatttagaa	gcaacctaaa	tgttctcaaa	cggaatgaatg
40681	gaaaaataaa	atgtagtaca	tatacccaat	ggcgtactat	ttaaccattta	aaaaaaatga
40741	gatcctgtca	tttgaacaaa	catgaatgga	actggaaaaa	atgttaagt	aaataataaa
40801	taagccaggc	ataaaaaggc	caacatcaca	tgttctcact	tatttgtaga	atctaaaaat
40861	caaatcaatt	gatctcatgg	acatagagag	ttgaaaagtg	gttaccagag	gctgggaagg
40921	ggagtagggg	gagggcaagg	gtggggatgg	ttaattgggt	ctaaaaataa	tgataatattt
40981	tgaataagac	aaactaatag	cacaatcagg	taattatagt	caataataat	gttatattttt
41041	aaaaataact	aatataattg	gattgttttt	aaactcaagg	ataaatgctt	gaggggagtg
41101	atccccattt	ctcatgatg	tgcttatctc	acggttgcatg	ccgttatcaa	aacacctcat
41161	gtacctcatg	gatattttaca	cctactacrt	acccaaaaaa	ttctaaaaat	aacttttttt
41221	aaatatgtgt	attataaaca	gcacaaaatt	tatcaaatat	ttgaaaacat	ttttgataac
41281	attttaaagt	gtgtacactc	aaaactgcaa	aaattgtcgt	gagaaaaatt	taaaaaatgg
41341	tgaataagtg	gactactgtg	taattggattg	gaagagctaa	aaattattaa	gatgttacctt
41401	ctcttcaaat	tgacttttta	attttttttt	attatacttt	aaagttctagg	gtacatgtgc
41461	acaaagtgcg	ggtttgttac	atatgtacat	gttggtgtgc	tgccccataa	actctgtgat
41521	ttacattagg	tatatctctc	aatgctatgc	ctcccccttc	cccccaacct	acacagggcc
41581	ccggtgtgtg	atgttccctt	tcctgtgtcc	aaagttgtct	tttgttcaat	tcccacattt
41641	cagtgagaac	atgtgtgtgt	tggttttttg	atttaattaa	atcccaataa	aaataatagt
41701	aggtctttact	gtagaactgt	gcaagctgat	tcaaagttta	catgtaaaaa	aaggacatag
41761	aatagtcaag	atagctttaa	aaagagggaa	gtctggaac	atgcaaacac	taagtttcaa
41821	gacatatata	gaagctataa	taatcgagag	agagtggtta	tggaataaaa	cgagatgtat
41881	agatcaataa	aaagagaaaa	aaagtccaaa	aaatagagttt	cacttacaat	gtcaagatat
41941	tttccaaaaa	gtaccaaggg	aattcagtag	ggaaaaataa	ggcttttcaa	aaaatgggtc
42001	taaaatttgt	agatattttac	atgtaaaaaa	aatgagacct	gattcatata	tgttacctct
42061	ataaaaaatt	actcaaaatg	tatcacagaa	taaatcaaaa	actaaaattt	aaagacttga
42121	ggagaaaaat	tttatgaccc	aggttttaag	agagatttct	tagatgtaac	tgtgaagaaa
42181	taattccata	aattaaaaca	tgaaaaaaat	agttgggctt	aactcatgtt	aaacatttct
42241	ctctcttgta	agacattgat	aagagataaa	acaaacaagt	caaagactgt	ataataattt
42301	tgaaaaatcg	atattggaat	tgtaaccaat	taaggaattg	tgaccaaaat	atcaaaaaat
42361	ctctcaaaac	acacaaaaat	gaacacagag	aactcaatct	aaaaactggg	taagaaattt
42421	aaacagttac	tctactagag	aaaatacatg	gatggaaaaa	aatgaataaa	aaagatactt

42481 aatattattt ataattaggg aaataaaaatt aaaaccacaa ttagataacca ctattttttcc  
 42541 atcagaatg ttaaaaattaa aaagattaat caaagtgttg gcaaaagctg ggaagcaactg  
 42601 gaacttcac acagtgctat ttgtaagtaa atatggtaca gccacttttg acaacagttt  
 42661 ataaattca tcaaaagtta aacataatc ttctctatga ctacagtaat ccagctggga  
 42721 tttatgcag agagataaaa gcatatgttc aaacaataat ttgtatatta atgtttatga  
 42781 agactttat ttaaacagct gaaacctaga agcaattcaa agtltcttca acaggtgagt  
 42841 ggattaaaaa aatgtggaat atttatgcaa tagaattgtt ttcaaccata taaaagaaag  
 42901 aactgttgac aaactcaatt acctgaataa atatacaatt attttgcaaa aggtgacca  
 42961 ggtaaaaaag aaggcatatg attttatttt atttacaata tatltctagaa aagtcacact  
 43021 aactctattg gacagaatgt ggttgttatt tgcattgagt aagacagaaga aggtgggga  
 43081 aagggtgggc aaagaatcac aaggacactt ttagaagtaa taaatacact tactattttc  
 43141 atcatagtgg tagtttcaca gtgtgtgac atatgccaaa acatcaattt gaggacttga  
 43201 aatatatgcc atttatttta tatctgttac acttccacaa gccctttaaaa gaggagaaga  
 43261 gagagcaaat ttgaagtgc agagatagaa tgaattgtct ttctgttctg ttctaccatg  
 43321 atgcccctcat tcaataagct gggtacattt cctcatgggt ttgtaggaca tataattgga  
 43381 atataatat atatatatac acacgctata tattatata acataataac atttataat  
 43441 atacataatg gtatttatac atatatatac ttaactglat acataataac atttataat  
 43501 acctctttaa aaaggccctc gagagacccc ttgttctctc caccatgtga ggttaaga  
 43561 aagaagtgc atctgtgagg aagtggggcc tcaccagtc ccaattttgt caaacactg  
 43621 attcttgatt tctcagctc cagaactgag agagtgaagt ttgttctctc ggaagagag  
 43681 agagaagaatg ttacacatg cctccatttt attactttgt tgactgttcc aaatcatga  
 43741 agttgaatct taaactatt cgttagctga aatggtttgt ttgtttgaa ttctcttaac  
 43801 tcaactatcc cttcctcaga aaactcact tggaaatttg tttaaatatg taaaagtaat  
 43861 cccaaaaggag agtgccctag tccagtcact catgattccc ttgttctctc atttcttaag gactgcatt  
 43921 agatcaaaa aataggagta gatatttcca acttaccagt attcttaatca ttggaactgt  
 43981 cccagagatct attcatagt ctgttctctc cagcataat ttcttaactg gctcactcc  
 44041 aattatgctt tctataattt tcttttaagt ctaacaaaga atgatactca aatatgtata  
 44101 ttgagcttct cattttcaac tgcctttctga acatatgtgc ttctgaattg tctgaattg  
 44161 aagatgcttt aaaaactatt atgacattcc ccttccaatt attatttttg ttattattga  
 44221 ggcagagctct cactctgtca tccaggctaa agtcagcgg tgcgactctg gctcactga  
 44281 acttccactc cctgggctca agcgattctc ctgctcagtt cctcctggta cctgggattg  
 44341 cagtcactgc ccaccaaacc tgactaattt ttgtattttt agtagaagc gggttttgct  
 44401 atgttgayca ggctggctct aaactcctga cctgaagtgc tgcgctctcc tctcttgcaa  
 44461 tattttatct ccgtcttact ttctatttt ttMgtatca ttcaaaaacc agaaagcata  
 44521 attaggtcaga gtgtggcaaa aaatagaaat gatcaaatca ttttgagggc Rtcaatgaa  
 44581 gctttcagtg aagttggttg aactgcgctc tgcaggaaga ggaagaagaa tctcttaagt  
 44641 aaacagaaga ttgcatgtct atggtctgctR tgccttgttc tccccaaaat atacaaaata  
 44701 tataatgtaa aatcctgcct ccaaagggtga ttgtattaga aagtgggact ttgtggaggt  
 44761 atttaggtca taaaggcaga ttctaatga atgggattag cacatttata aaaaggctcc  
 44821 tgagaacccc ttgtccctt ccaactatgt agtacacagc aaaaagggtg catctgtgaa  
 44881 gaagtgggct ctcaccagac accaaatttg ccagcactt gatcttgtag tttccaggct  
 44941 ccagaactct taaaataaaa ttgtgttgt ttataaacta tccaggttat ggttatttgt  
 45001 tatagacagt cagatgaact aagacatgca ttcaaaaaca aagaaaactc cataaaccaa  
 45061 tgcataaact aagaacagtg gcatgtgcct tggggatctg caagaaagt cttatgtgt  
 45121 ttgcacagca tagagactgt atgtagaagg aaggtcacaca gacaggaaga ggggtccaga  
 45181 gaggagtaat ggtattttt tgaattttaa aataattcRg tacagacaga gggatattg  
 45241 tgtttatcta caactaYta taaatctaca catctcaccc cttgagtatg ttgttaact  
 45301 ttgagaggaa ttttctaat attgcagttc cccaggaaac actctcgaat cctaagatt  
 45361 catgtaggag gttactgtgt agcgtttctca agaaccagag gtgagagaag ggaagaaag  
 45421 ctctatctgat cgagtggcaa cagataagtc agctaaggaa gtagactccg ggcctgggaat  
 45481 gactcttcag agttatgtta attaaagcaa gtgggctgga ccagaccttt gtgctgtcca  
 45541 catcaagcaa tcaattggctt ccagctgcct ccaagaggag gctgtgtttt ttgtctgtta  
 45601 tattattgag ctatgctgca gcttggtttc tccagaagac aaaaactgaa acagagttag  
 45661 gactgcagct taattttgct atgtggcaga cactgagacc tacttaggac ttgcacattc  
 45721 cattaggaaa tcaaatctgt gagagaaagt aggcagaagc ttgcttaggc aaggggagga  
 45781 acagaccacc atgtggagct gacaaagtct ctcttagccc agtagggagt tccagagaca  
 45841 agatttatga ttacaggagt cccgtgatgt gtgaaatata ctaggccatt gtacaccac  
 45901 ttgttgcact cctctggccc aggccacatc gagaagaagt ttacgttgtt gcaaaaagt  
 45961 agggagaagc aaactaaaaa agccaaaagc tggagagatg cagcgaacca cactctctgc  
 46021 attcagacag taagggcttt cttgaagaaa gatctgagct gatctgagct ttgttttacc  
 46081 atgtctgcct catgacttca tgactttctt cttctctctc cttttttttt tttttgagac  
 46141 ggaagtctcact ttgtgtcaac caggctggag tgcatttgtg caaltctcag tctctctcact  
 46201 ctccacctcc cagatccaat cgattctcct gcctcaggct cccgagtage ttgggtgctct  
 46261 cgtgatctgc taagatctat ggtgtagtta tgggtgctct cctttttgtg gaactctctc



46321 gtcocgggttcg tgagaaacctaa caaaaatattc tgttatttga ctaaccatttc cttgaagatttt  
46381 cttccaggcgg tgggattttta tctttttttaa gtttcccttga atatacaagg atatacaagg  
46441 caaatgagaca tataatgaag tcaactttttt agatgtacaa ataccacaca cactgtgtacc  
46501 tctgaggttta ttgaaactgc tgaatagagt aactttttat ttccaatttc ttgctaattg  
46561 agagagaaga ttgcctctgt gttaattgttg gatttttaaaa tatatccctt ttaagcacta  
46621 ttaagcatgtt tttaggttac aaagttatttt tatctcatct ttccaatttc aaagaaaaaa  
46681 atattattaca gaaatgtctc aaactttttct ttcttcaggt tcaagtctcat tgaaaaagaa  
46741 gatgtctaca tcaattgagaa gaaaaagaag cccaattgtg agtagcaca atataacagg  
46801 agtactcttgc aaggccacaga cactgcagga ataaatttga ggctgcattt ctgtctttag  
46861 gactcagctga aatgggacat caaagccatg caagccaaca gcataattga aggaacagtg  
46921 ttgaaaaaac aatatgaaca agttttttata ggaatcaaag aaagaaagaa ttcgacttga  
46981 tctgaggagt acaagaatat tccgaatttg gggaataaat aggaagagat aacagaaaag  
47041 ccaattgaat cagaaccagc aggcaggaag tatgaacca tctagcagcc ttttctgtaa  
47101 ggacaaaaga tcaattaaata ggcaaatgtc agaaagccag ggatagtatg gttttgattt  
47161 ctttaatttcc tctttgttga gtttcccttt gtcttagaaa ttaattcagga ctttaagacct  
47221 actgtgtgat ggacagcag gaatatcagg cctattaatc Rcaatcaagt gtaattgtct  
47281 tgtatttaata aatcatgtttt aatgtagaat aatttctgaa aataattgcca cagatcaata  
47341 aataaaacttc gagaanaagc tgaagttccc agggctcat tctgagaactt agggccagga  
47401 gagttttagac acttctaagta gaagacaccc aggcgaagact ttttgggata aggacaattt  
47461 tacaagtttca tgccttcttg aataaagctt agttctaata actctagaaa aacacaaaaa  
47521 taaatttagct atactgtatgg tttttggaga acaatttttcc ccagattttaa aaagatgat  
47581 ggatgtggctt aaatcaactgt ccttaacctc tttatagaac cagaggtttc cagccttaag  
47641 attttaatttc cagggttagg atgcaagaag tgaacttact tttaggttcta ccagggaagc  
47701 aaatggcaca catcaaaaaa aactcgtttg cgcagtttcat agattttcaa cagtttata  
47761 aaatgtacag tcatcagctt ttttcttatt gcaataagaa actattttaa tgcataaaaa  
47821 ttgtatattt ccagcttcaa aatgaatat gtcactcaac agaatttaag cagttaactac  
47881 agaaagagta atcaataaat aggccttctcc tcttactacc cttagaagatg gggagatgtt  
47941 attttagtcaa ttgtttttct aaaaattatg actgttagaa cctagagctta aatttatata  
48001 gaaaaagttc aaggaaatga tgaacttatt ctttaagaag gtgatttaag cctctgtgatt  
48061 acaggaggtt taagatgttg tctttgttgt aatgcagtgt ttttaactta ggcccatcgt  
48121 gaacatacat tttccctcat gaacaaaatc accagccaac caaaacctct attacttcta  
48181 ttttctctga aggacttctc catctccaaa ttgagtgtat ccatgtcagc cttcataaagg  
48241 taccaccgaa atgtggcagt ctcatlttga gacacacatt acaaaagagt ttcagccaaa  
48301 atagtttgaa tcaagaagtg gtctttgaa atgtattcag acttaatacat cactaagaa  
48361 gagaaacaaa tgaggcatg gagaccttcc atccaagtgt caaataggaa aagtctgtca  
48421 tgtaccatgg catgtacagg aaagcagtg tttttcaaga atctattcag aaMgactatc  
48481 agatttttga aacaaagtct gcatttcatg atgccccaaa atctagtgtga aaacataata  
48541 ttcaaatatt gttaaccttt tttttgtata tttatttgtt atactagacta gaattattca  
48601 cttatttatt gaagtattgt tcttatagga agctaggcac caaccagcag ttgctaataa  
48661 tatactataa tttagtaagt aattcaattg aataaaattt atgacaaatt taacaaaact  
48721 gtttaattgaa atgcatactc acgctgcagc gtccacttaa tcttttgtcc Rccagtgcct  
48781 atgcatgtat tagtatgcat caaatttttg agcagtacac aagtgaatc tctgagactt  
48841 cccccacga aaaaatgat gatataaac agttatgatc agatccccag agtgttccct  
48901 taaactgtat gggggccaa ttgtaactc ttgtgtctt acactgtat taacttcca  
48961 gtatctattt ccccatattc cttataaata aaacctagat tttagttggg aacgaagggt  
49021 tgcctcagta aaactcattt ctctaaccaa tgtgatgcca gtgctttgcc ttttctgtct  
49081 catccctctg aataaaaaag gcaaaagctc ttttgccttt tatttttact tcttctgtct  
49141 tgatagctag gtataatc atgttttgtt gtagtggatg gtagtggatg gtagtggatg  
49201 gtaggtgtga gaaagttaga aagaggttaa atccttgata acactgcaga atctgtgctt  
49261 ctatcatlag cttaatttca ggcttctcac ttagtgagaa aatatagacta atttggtaa  
49321 ctattgtctt cattttgctt aagtcaaaaa ttttcttctc ctgtgtgtcc ttttgacaa  
49381 atcttttaag ttataaatat atagaaatat cacttttctc attttattctt ttaacagtat  
49441 taccttgggt ccagaaactt gattttgagt aacaaattgt taattgttga aaactttctt  
49501 gaaattttctt gctcttaaca agagccagaa agagacclgt tccatattct atttactatt  
49561 gagccagatt gtcaatttct ccaggccctaa tccattactt tgtgatatt gatattcag  
49621 tttatgaaaa ttgattatgc acaagtttct ccttgcaaat cagatgtcaa tcttctgaa  
49681 ggtccttccac ccaatttatg gcagccaccc agtcaattcac aaccacggca tcttgtttt  
49741 attttttgta tggccctgat aaacctctga tacttttcta atcttacttt ctgtattt  
49801 acgatgtgtt gcttttgttt acttttagta tcttccccca agtgcagga atgtgagga  
49861 gaaattgtgt gcttttatct gaaatatact cttgggttat agagagcaag taggcatatt  
49921 tctgtgaatga ctgaatgact gaatcagtaa ctaaggaaatg caaagaattt tttagatatg  
49981 tgaactcagct gaggaattt caaagacaaa taaggaaatg agagacttcc ttaaaaaat  
50041 gaaaaaatga agttataaga atgtctttaa aaaaaggcca agtgtgtgtg ctcacattgc  
50101 taattccagc actttgggag gccaaaggag gaggaattt tgaagtcag agttttagag

50161 cagcctggccc aacctggcga aaccccatat ctactaaaaa tacaaaaaatt agccagacgt  
50221 ggtgttgggl gctgttaatc ccagctactc aggagtctga gacaggagaa ccttgtagac  
50281 ccaggagggg gagggttcag tgagccaaga tgcataccact gcaactccagc ctgggcaaca  
50341 agcagacttc tgcctctaaa aaccaaaaaa aggatattaa aaggacagag ctacagctcag  
50401 aaccacaatg agataagcca caacttggat ttttagtag ctacagagac caagactctc  
50461 atattgattc agcagcaga gagagcaaat ccatagtga aggaaggagg ctgagacaaa  
50521 gcatgttgta gcataacata tctgtagtt taaccagtaa aaaaagcctg gtgtctcttt  
50581 gagaacattt gctagtgtgt tgaaaaatg ggaactctg ctcaaatggt ttgtgttttag  
50641 gattggctat gctgagttat tagtgtagtg caaaagtta tgcagttttt gctagtgttg  
50701 aatttgcatt tgataattgg aataatactt taataaatg tgcattagtt atcatacttt  
50761 taattggpcc ttctcaactt ctattttttc actaatgaca ttaacttgcgt tttattttgt  
50821 gtttatttta gacaaaggaa acaattgttag ataaaaagca aattctcagc acttttctat  
50881 ttgagttcaa aatgggtcgt aaagttaaga gcaaaagaca ttcatcatat taacaaagca  
50941 ttgtggccag gaactgctaa cgaactgaca gtgcagtggt ggttccaagat gttttgcaa  
51001 ggagagcaga ccttgaaga tgaggagcgt agtggccggt catcagaagt tgacagtgac  
51061 caattgagag caatcatcaa agctgatgct cttaacaagt aattctcagc tctgtgttag  
51121 gtcattgtgc accatttctac ggttgttcag catttgaagc aaatcagaaa ggtgaaaaag  
51181 ctgcagattt ggtgtgtctg taagctgacc aaaaatgaaa aattctgtct ttttgaagt  
51241 tctctctctc ttattctaca caacaataac gaacctttt tcaagtgtat tctgtgtgc  
51301 aaggaaaaat agatgttata tgataaatg caatgaccag ctcaagtggt ggccgagaaa  
51361 gacactccaa ggcacttccc aaagccaaac ttgaccaaat tcaagtcagt gctactgttt  
51421 gatggtctgc tgcggtgtgc attcaactca gctttctgaa tccagtgaaa accattactt  
51481 cttagaagat cgctcagcgc atcaatgaga tgacacaaaa actgcaaacac ctgcagttg  
51541 cattagctaa cagaaaggcc ccaattctct acacagccca actctcagc accaccacaa  
51601 tactttccaa gttgaacaaa ttgaggtggg aagtttttgc ttatccacc atattccact  
51661 gaccccttgc aaccacaacta ccacttcttc aagcatctca acactttttt gcagggaaaa  
51721 tgcttccaca accagcagga tgacagaaa gctttccaag ggttctgtgc attccaaaagc  
51781 attgattttt tgctacaggt aataacaaaa ctattttctc ctgtgcacaa ctgtgtgtg  
51841 tgtagtgttt cctattttga ttaataaaga tgtatttgag cctacttata atgaattaaa  
51901 attcagattt caaaactgca attactttta cccaaactta atagtattla cctagctttg  
51961 ctgagcctca agttctctat ttgataaatg gggatgagaa tgcattacct cttaaatgtt  
52021 aatcagattt aaaaagacga gtatgaaagt aatttgagaa gcatgagtaa ttatgaaat  
52081 ggctaaccag tcttgagttc ttattatgtg tcaggcactg gagtcatcac ttccacaga  
52141 tgatctcttt taactctcag acaacccoa aggaaaaagac aatatccagg ttccaactg  
52201 aaggggagag agccagctag gagcagaact gggctgggga ggggtttggg ttttgggtga  
52261 tggaaagaaa gctcttggga aaacatttct tcttggacc acctcatcca acattgtctc  
52321 tgatgatcca cagttccacg acttatttca cacttaactc agggagatga tatccaaacg  
52381 tggccaccag gtgagacata tatatacatc ccaatttccac caagcttcca aggatctact  
52441 ttggcaaaat acattcaaca acttggggtta attctgtaac ttccgtggag aggaagtga  
52501 gctcttccaa atttactttt tgcatccaat ttgtcacaag ccaagtgtat ttttaaagt  
52561 acattttctt aactgccatc acatgcaaat ttctaaggaa attagaatgc tggcaaaat  
52621 ctctccagaaa ttgaggtagt aactcttgca cagaaagaa actgtaccata agtccacaaa  
52681 tcttgaatag aattccaaat attcagatca tgatttacc aagcagaatg cattaagaat  
52741 taggagcata gttctcaggc tgacaaatct gtttactgcc ttggtgtcac atattgtct  
52801 gtaaaactgg ggaattttat ttattcacca tgtaaaaaca tttttctcaa gttgtaaaat  
52861 gaacacaaa aaacgtattac tcaagtgtca tctgtgaca taagttagat gatataatg  
52921 aagcaaaagg ttgtgtgtgt ctacagggaca gtactgaaa ttgtaaatat tttctttttt  
52981 actactaatt ttgtaagtct ttcagatttc tcaattgtta tcaattgtta ttttattatg  
53041 tgctactact gacaaaattt aggtttttat gtaatttaag tattggaaaa tacaataagc  
53101 atattactct ttgataatt atccttttag ccaagttag actcactcat agcccaattt  
53161 acataatgct aaatgctgaa attagtttac ataaaagtta tggggccggg ggcagcagc  
53221 cacccttgta atccaagcac ttggggagcg cgaggtgggc ggggtgggcg ccaactctac  
53281 cgagacccat cggcctaaca cggtgaaact gctgtagggc gctgtagggc tgcactgagc  
53341 cagcagcagc ttggggcgcc ttgtagccca gctgtagggc gctgtagggc tgcactgagc  
53401 cgtgaacccg ggaggcgagg cttgtaggta gctgtagggc gctgtagggc tgcactgagc  
53461 ggccagacag ctgagactcca tctcaaaaaa aatagtaaaa tagtctgata tgcactgagc  
53521 atattttgtg ttttagtatca gttctcaaaa ttgagatagt acatgtgata tgcactgagc  
53581 aacccaagcc aaacctacta ttaataacta tcaatgaatt tttgtttatg gattacttta  
53641 tacattttca tagttctaat cagcatggga attttcaaaa tcaatcttat gattacttta  
53701 attgagattt ttttttagtt ctcaagaagg attttctctc ctttaacaga aattttacat  
53761 ctcatgtttt tgaatcaata tatgccaata gattttctct ctttaacaga aattttacat  
53821 aatatattat tctctcaatt ttgtattctc tatgctgtgc gaaaataatg agaaataat  
53881 taataataat gttattctat ggaataatga ttaactctgc tttctgtgac tttctgtgac  
53941 gtaattttta gattggaatt ggcctctat tttaactctg agattttttt gctctctctt

54001 aagagacag aagactccta attggaata aaaattgtag cctagatact tatgcaaaat  
 54061 ttttaatagga ctgcgctgtg taagtctcgaa attcactcc agactatcca gattcaattta  
 54121 tattttttatc tatttagaat ctatgtttgtt atatttttttg tagttgtaat catcgctgtt  
 54181 tcccacgaat gaatgcaaac tctttttagt ttggaccatt taattgaact ttggaacatt  
 54241 caccccaacg tccaaggtca tctgtaacat aacagatctc gcgcgtccat taaagtgggt  
 54301 gaatgaact ctgagggggt aaattatgcc tcaagaaggt tgggtattca ggcctcagcc  
 54361 accacacagg gctttttgtt tctgtattga ttgtctcttc catgtgcttg aggcctccat  
 54421 ggcctctcag gctattgtct gcgtctttac tcaatttatg ccagagaagg taagaagaga  
 54481 tatcttgaaa atttggaata aaattatctg ggcctgtgta ttggttcatt ctccactcgc  
 54541 tctggaataa tacttgagac tgggtaattg ataaagaaaa taagtttaat gctgcactgg  
 54601 tctgcagcgt gtacaggaag catagcagct tctgcttcgc gggaggcctc agggagcctt  
 54661 tactctagcg ggaagcaca gcgagagcat ggcctggaga ggaaggagg agagagagt  
 54721 gggagggtgc acaccctttg tttttgagac ggaatttcga agtttcactc ttttccagg  
 54781 ctggagggtc atgctgctat ctgcgctcac tacaacctct gcctccagc ttcaagtgt  
 54841 tctctactct cagcctcccc agtagctggg attacaggcc cccaccacca cgcctgctca  
 54901 atttttgtat ttttagtaga gatgggggtt caccatgttg gtaggctgg tctcaaat  
 54961 ctgacctcag gagatccatc cactcagccg tcccaagtg ctgggattac tggcgtgag  
 55021 acacactttt aaacaagcag atctcttgag aactccatca tgcgaacacg accatgggt  
 55081 tgggtctaaa coattcctga gggaccaccc ccaggaccca atcacctccc accagccccc  
 55141 atttccaata ttggggatta caattcaaca tgagatttgg acagggacag agactccaa  
 55201 catgtcagcc tagttctact aatcctgccc tctttttatt ggaattctcc aggtctctgc  
 55261 tgcctctcgc agcagagaaa gacctgttct aYaataatcc agcactgcag cgtgcctcga  
 55321 tcatctttcc aacacgtggg tcaacttagca agcctatttg gcctcacaca tctcaact  
 55381 agctcttact agtaactgaa atggcacttt gcctctgctt cctctccctt tatttttacc  
 55441 aaattgtcac aaatccaggt gaagaaaaa ggtgctcttt gttgagcttc cttggaaaaa  
 55501 ctgattctat acttttattt ctctcttgg taattcttat tctgtgagg gttgatacc  
 55561 ttcagaggtt atagctgaa atcctaacc ccagggtgat gatgtaagg ggaagggtct  
 55621 ttggaaggtt atgagatcat gggagtaaat tctgtcttt taattcttat tctgtatgc  
 55681 atttttgtat agcttcccaa ataatgaaac ccaatagcta ctgtgtaatt agtagtatac  
 55741 ttaacctcca ttttattcca ataatgttaa tcaattgact cctctctatt agacctcag  
 55801 ctctacgtgc acattttatc tctcccatc acctctcacc tttccatcag agcgtctcta  
 55861 attttacatt ataatgccta ttaacgttta ttttctgttg tataatcata atttaaat  
 55921 tatgctatct ctgtggagggt attctaagac ctgaaacta gtaaacagca gttacagat  
 55981 aatgactaca taaatattca atgaaggccc aactcagctg ctactatct atttcttca  
 56041 ctttatcccc atatacccca ctcccatagc aatcaacgtg taatgatcct aagcatcctg  
 56101 gtcaaatgat tctccataat tatacttcaa tattgtctca aaatcatgccc ttatttattg  
 56161 tatctcacac ttggacacta actttataga cagggttagc tgtttatttt cctagagttc  
 56221 atatttgtct tccaggcatt tgtatttatc ctactgaga gaagaagcat atttattctt  
 56281 cactgtgtca ctactgtaaa cctgctcatt tgtcatcga ttttataatt ggaatgccc  
 56341 tteactcttc tgagaacctc ctgacgctc actaacttca agtttccatt tccatttgt  
 56401 tactttctgc gccactgtc cacttgttat gctgagatt cttttgactg acttcttagg  
 56461 tcagctcagc tgggttttYt acctgttatt ctctgacatt caaacactgt cactatgca  
 56521 tctctatgac atttaccagc tatttactat gtacaaggca tagtatata tactatgaa  
 56581 attcaagcga aatcaaaagt ttgtctgcac caaaagagt aataggttaa tagaattga  
 56641 cacataaata taatcataac ataaaaacta caaagagat aacagagaaa tagaattga  
 56701 caaagaattt agaatgaaa aatagaaat atgcataga tgaanaMtt agagagcat  
 56761 cagatgtgaa ggggtgaagg ggagataggg catggtgaga gagagctcaag gaaggtcaga  
 56821 gaatgcatga gaagaaaga agcaatctca cctgtttaa caacggttgt ataggaaat  
 56881 agacactttc atagtgaaca aaattgtaca ctgctattct acatagtat aagggatctg  
 56941 gccaggccaa tcaggccaaga gaaagaaata aagggatctc gtatatttag aaaaacccat  
 57001 aaattatctt tgtttgcaga ttacatgatt agtaatttag caggatacca aatcaatgtg  
 57061 caaaaacttg ttaagctgag aagaactctc aatagacaaa cagccaaaat ctgagcaaac  
 57121 tcccaattcac aattgtcata aagagataaa aataccttag aatcacactt gcaagggtg  
 57241 tgaaggacct cttcaaggac aactcaaaaac cattgttagg ggaagtgaag gaggaacca  
 57301 acaaatggaa gaacattcca tgcctatgga taggaagaat caatatcggg aataggcca  
 57361 tgcgtcccaa agtaatttat agatttaagt ctatttccat caagctacca ttgactttc  
 57421 tcagagaaat agaaaaaaa aaaactactt tgaatttcat atggaacaa catcaaaaaa  
 57481 agagctctga cagccaaagc aatcctaagc aaaaagaa caagctggag catcacata  
 57541 cctgacttcc aacaaaaaga gcatggtcat ggtaccaaaa cagatatata gccaatgga  
 57601 atagaacaga ggtctcagaa ataacaccac acatctacaa ccatctgac ttgcaaaaac  
 57661 ctgacaaaaa caaaaactg ggaacaaat cccattttaa taagtgtgtg tgggaaact  
 57721 ggctagcgtt attcagaaaa ctgaaactgg accccttctc tacaccttat acgaaaaat  
 57781 attcaagagt gattaaagat ttaaacataa gacctaaaa cactgagaaga

57841 acctggggcaa taccatttcag gacataggca tgggcaagga cttcatgact aaaaacccaa  
57901 aagcaatttgc aacaaaagca aaaaatgaga aatacaatct cttctggaga cttctggaga  
57961 gcaaaaagaa ctaatacag agtgaacag caagctacag aatggggaga aatttttga  
58021 atgcatcttt ctgacaagag gctaatactt agaacctaca aggaacttaa aacagtttag  
58081 aagaaaataa caaacacccc catcaaaaag tgggtgaagg atatgagag acacttttca  
58141 aagagagaca ttatgtgac caacaaacac ttgaaaaaaa gctcatcatc acttggcatt  
58201 aagaataatc aatatcaaac cagactgaga tatcatctca tgcagagttag aatggcaatc  
58261 attaaaagtg cggagaccaa cagatgctgt agagagtgt gaaacatttt aacactttta  
58321 cactgttgtt gggagtgtaa attagtctca ccaattgtga acacagtggt tcatattctc  
58381 aaggatctag agctagaagt accattttgc ccagcaatcc cactactggg gcgataccca  
58441 aaggattata aatcatctca ctataaagac acatgcacac gtatgtttat tgcggcactg  
58501 tctcaaatag caaagacttg gaacaaaccc agatgccat cagtgttaga ctgataaag  
58561 aaaaattggt aggtatatac cgtggaatgc tatgcagcca taaaaagaa tgagtctgtg  
58621 tctcttgcag gaacacaggt gaagctggaa accatcttc tcagcaaac aacacaggga  
58681 cagaaaaata aacacccgat gtctcactc acaagtgaga gttgacaat gagacatat  
58741 gggcacaagg aggggaacat cacacactgt agcccgtag ggggtgggg cagtgttaga  
58801 gtagatgact aggagaaata cctaattgat atgacaggtt gatgggtgta gcaaacatt  
58861 atggcactat ttaacctatg taacaaacct gcaattctt cactgtatc ctgaaactta  
58921 aagtatata ataataataa aaaaatccct ttgccttctg ccatgattgt gaggccctcc  
58981 cagacatgtg gaattgtaa tctaataaac ttctttttct ttaataattt aaaaaattga  
59041 tacaagtact ttagtctaaa gaaaaactaa aaccaatttc atccatttag aaaaaataa  
59101 aactgaatat tgaaacccat gggacactgc caagatcatt ctcaagaaac aatccatgac  
59161 aataaattgt gtattattac gggagcaata aacctgggag aaaaaaattg actaagcaat  
59221 aacccaagaa cactgaaaaa atatctaaaa aataaattct ctcaagaaac aatccatgac  
59281 gatgaatata aactttaaat ataattgagt agacatctaa ttataaagttt gaattcaaac  
59341 attaaaaccaa ggtgtgtttt ttaggaaagt taatggaaata gatgaattgt ttggccactt  
59401 ttttttaaga caaagaaaaa tcatatatgt gtgaaagata cagttaatat aacacatat  
59461 attattataa aagtagtgtt atataaata cacatagaaa acttgaacac tcgataaata  
59521 gatgtaattg ttgatatta tatattatta aagttgccta tatatcagaa taactaata  
59581 actcaaaaat aaatgttttt agtgcgaagg aacacccctc gccctcccaa aagggtcaag  
59641 aaaaatcaagt atttctatat aaagaatatt gctttaaagt tcttaggattt caggttaatga  
59701 tgatgcagac aggaaggtgt cagactaatc taacttctca ctaaaaccca aattctaaaa  
59761 taaaaataac agataaatca aatctaagtg agcaggaggc ctgtttaaat ttattctaaa  
59821 aatgaataatt attatttga cagggataat aaaaaggaaa agggagataa agaatcaaa  
59881 gatgattact tggagtgtaa tttgggatac ataataagaa gtaacattat tcaattgagt  
59941 aagaagaagca gcttcgtgta gattgaaata tttgtcttga atttagatat aaagtttgtat  
60001 atatatatac aaatttggag atataattac agtataaact gaataagata aagcattagc  
60061 aaaaatcaact agagagagcc ttcaagaattc taacagaaca aattatctct attaacttta  
60121 caaaacatat ttgtaattt gcattaacca tctatttcta tatatataga caaataacct  
60181 caaaatctat tcagcagttt aaaaactcaag aataaactat aattatctct tgcatagttt  
60241 ctgagggtaca ggcacctggg agaattctag cttgtgctatt ttggccagaa tctctctatg  
60301 atatttgcag caggtttgtt gccaggactg ctttagccat cgcaaaatatt tctccgggct  
60361 aggggtctcta cttccaagct caactgtggc ttttcttga gggctcagtt cctcaaccaa  
60421 tggacctgtc ctatagtcgc ctcaaaaaat ttaagtttgt tatccccagc ataaaggac  
60481 taagagaagtg tgagagagat atacacaaag tggaaagtaa atcttataac ctaatttgg  
60541 aagtgataacc caatcatttc ttctatatct tatctttag aggcaagtta ctaaatcag  
60601 ccaactcaacc agtaggaaaK ttaagctctt actgtcctaa aagagatcaa caaataatgc  
60661 caacgtatct ttaaaacctt cacaatatMcc aaccacttg aagactctag aatttgttat  
60721 gagaaccaag acaaaactat taggaattcta cataaacatg aaggcgtag tgatgtcctg  
60781 gaagtcaagt gatttgaatg agccactgtg ttttaagata gtggaaagta aataaggaca  
60841 ctgagtgccc actatatacc tggcactgta ctagggtgtt tatgtgcgaa cctctctttt  
60901 cttgcatata atacaacca tgaacttcca gtgccttcca cactgtgag atcctgtttt  
60961 ctgaatcaga taggataaat ttgtgttttg aaggacagag tctcaagata gttaggaatt  
61021 aacacacata gaaacttatct ttctcacaca taaaaatMtg gaagtaggca gactaacgta  
61081 atggggagca tggctctgct ccatgaaatc ttggggcatt ccagcccaYg ttctcatatt  
61141 tgaatgccca cagtgtcatg atgcagaaca cgtctagagc tccactctttc aagaagcaca  
61201 atggaggaaa gattaaagac aaaggagcaa aggacaggtt ccagctgtct tttaggagc  
61261 tcaacaggaa attaccaaac acttctactt acctgtcaca acttttccat ggaactgaga  
61321 attctcttat tctgttagta ataaagccca gataaaaatc aggggcttta ttaacttga  
61381 aagatagata tatgggggta aatagtagtc tttagatagg tatcatatgg aataaccaga  
61441 agcaattctaa agtcaggagc ctctgaaaaa attatagcc tttagtatcc tccagcatt  
61501 gtgaaatctg ttagtctctc aagctcaagt ccttggacta attttgacta ttgttatgat  
61561 agtttatatt cttagtYgaa gtacacaaat cttatttcat tctaccttaa cagctcacat  
61621 tatcatgaat gagaagaaat gtcacctgt attttaaaaa tttttggcca atttagcagt

61681	cttgggttctc	agatgtccca	tctagtaata	atatcttctt	attggctttt	gctattttgt
61741	gggctagaga	aaccacttgg	aaggaccata	aaagtcccca	gttttatgtt	ggtactaaat
61801	atatctcttg	agccattata	tgtccctacc	aaatttttta	tcaatcaatc	aatgtacaaa
61861	ttatgttatg	gtgtgattat	ctgtgtatgt	ctcaaatccc	acaaataatt	attagtaaaa
61921	aaggactaga	gccacagttg	gagaaattga	attattttat	Rctaataatc	ttaagcaaaa
61981	gggaattgcca	tttttacctc	ctcacacata	aaaaaatgca	gtataaagt	aaagtgtttt
62041	ggcctctaga	aaacaccata	tgagtttttg	ggctcagaaa	attatacccc	aaagtgaagg
62101	ctctcagaag	agactcagaa	gcaaaagttc	tctgaacttc	tgcccgctgt	ctctcgcac
62161	ctcactctct	cccccaatgc	aagtcataga	aacttgtaatt	cttccctccc	aaagtgggtc
62221	atagaaaacc	aacccttttt	ccccaaagcc	agccataaaa	gttcaaaaca	taccoccaa
62281	cttcccccct	ctttcttgtt	aaaagatggc	cataaagaaa	ttctctgacc	tgacttgttt
62341	gaactgtagtg	cgtaagacc	cctgagtcca	gaaaacactt	cataccccg	aggaaggaat
62401	gcgttactga	gaggccaaga	agcatctgaa	cagacagacc	atgctgagtt	ttcccaatta
62461	gtctattccc	atcaggtcat	acccttttgt	ccaccatat	ttctacacgg	ttttctatc
62521	ttcatcaaat	ctaagcataa	aaatagacag	ttgtcactgt	atctttgagt	ctccactttt
62581	aaggtcccca	tgacacataa	aactataatt	aaaaaaatt	gttatgtctt	tctctgttca
62641	atttcttttt	tggtatagga	gtgtcagctg	tgacccttat	gatagggag	aaagggatca
62701	ctctctttgt	ggccctacaa	gagaaaaaca	caggtcaaa	gaggtcgaga	ctaatctttc
62761	gaactcttaag	ctcagactta	tggttaccac	caattcttat	ttgtcccat	ttcaacacca
62821	cttcaaaaat	attttatcct	caactgagct	cagaaagatt	aaatatcatg	ttccacttat
62881	catacctaat	aaataccagg	attgggattg	taaatctatg	tcttcttaac	gtcaaaaac
62941	atcttccata	gttttccaac	acaaggagac	ttgaatgcaa	tttcttttat	Mgttgcaatt
63001	ctttgttgta	tataatttaa	gagcattgtc	attgatctta	gggtgatttt	ttcttataaa
63061	gttttttgta	aatttctatt	tattgacatt	tgataaatt	tattgcaatt	tattgcaatt
63121	aaacttttgt	ttaaaaagta	tattgtttgt	tggtgtttgt	tcatttttaa	atgttaattc
63181	aacatattaa	aattggggatt	atctagataa	ttactggaa	tctctgata	gtgacctagc
63241	ttgtgtccca	catatagcta	ttctcaatc	caaaagagag	aggggaatt	ggacactgaa
63301	ttattctgcc	agctgactcc	gcactccttg	gtgactgata	attttctata	Mgttgcaatt
63361	ggcagtgcta	gcaagtaagg	attaatcat	tttaactcct	cctgtatttg	ttctttttt
63421	tttttgtagc	ttaaactatt	tccagtcttc	aagactttct	attgtctagt	tgtaattttt
63481	acaaatcata	ttttcacaga	ttctatggat	ctttctagag	tttttttaag	tgctctactg
63541	tagttttatt	tcaaatatat	tttaacatgg	tgacagtctc	ttcaactgtg	tttccctttt
63601	attgcaattg	cttcagagct	aatctcctga	agtacaattt	gatcatgttt	ttcagataaa
63661	gcttgaaaag	gcgtgaagtc	atatcagacc	ttttagtaat	cttaacagtt	ccgaatcat
63721	taaggcgaag	gagttagaaa	aaaacatttt	caatgggata	tttatgtttc	atgactttag
63781	agagaaattt	aaaaacaaga	aatttcaact	tttatcaaac	cctctcaata	ttgtttttta
63841	ttctcatgat	aactttttac	ttattttta	tcaattttta	aatttttata	ttttcatagg
63901	aatttttaga	gtactgtaga	acacaYtgaa	gttctgact	tctttttaca	gggtgaatta
63961	ctgagattta	ggttctatgg	ttggcccaag	tttcataaat	aatataatgga	aaataaggac
64021	ctaaagccta	tagagtccaa	tttctatgta	cagaatctct	ctgtaacttt	tgctctcttt
64081	gtacacattat	agggatcagg	agatgtcagg	attgtcfaat	acacagtagt	ttttcaatat
64141	aaataatgat	tagatataca	tagaaaatatt	tttaagcaat	tcagatctgt	cccaactctc
64201	ctactccctg	ttgctaccat	ctaagatata	tctactgcct	ttgaattctg	cagagtggga
64261	ctcagacatt	cataactagg	gcactctgag	acagctggaa	cttccctgaa	gacacagctc
64321	gtactcagct	accattggca	agatgaattt	tctcatagca	tcatacgtta	tggtttatcg
64381	tatttccca	cccatctccc	tagggccagg	gtcatatttc	taaatccccc	tttttatcta
64441	ataatgtatt	ataaacttga	aatttgcata	gtatagatct	taagtttct	caccacacat
64501	acacacacaa	aagtgtaaat	atgtgtgata	tgctgtatta	gctgtttttt	acgctgctga
64561	taaaagacata	cctgagatgg	caatttgcaa	aagaaagagg	tttaattggac	tcacagttcc
64621	acgtggctaa	ggaggccctca	caataatggc	agaaggtgaa	aggcacttct	catcgggttg
64681	cagacaagag	aagagagcct	gtgcaaggga	acccctcttt	ataaaacat	cagatctcat
64741	gaggcttatt	cactattaca	agaacagcac	aggaagagcc	tgcccccat	atttcaattac
64801	cttcccaactg	gtccctccca	tgacatgtgg	gaatttgtaa	agccacaaat	caagggtgaca
64861	tttgggtggg	gacagagcca	agccatatca	ttccaccctg	ccccctccca	aatttcatgt
64921	cttctacattt	caaaattaat	catgctctcc	ccacactccc	taaaagtcta	acctattcat
64981	accataactc	aaaaaccag	aatccaaagt	ctcatttgag	acaaggccta	ttccctccac
65041	ctatgagcct	gtaaaaatcaa	aagaaagcta	attacttctc	agatatataa	tatctatata
65101	ttctatagata	tacataaatat	ctatatatct	gtatagatac	ataatatact	tatatctgta
65161	tagatacatata	atatctatgat	atctgtatag	atacataata	tctatatact	tgatagataa
65221	cataaatatct	atatatctgt	atagatacat	aatatctata	tactgtgata	gatacataat
65281	atctatataat	ctgtatagat	acataatctc	tatatatctg	tatagataca	taaatctcat
65341	atctctgtat	agatacatata	tatctatata	tctgtataga	tacataaat	cttatatctc
65401	gtatagataac	ataatctcta	tatatctgta	tagatacatata	atatctatct	atctgtatag
65461	atacataata	tctatataatc	tgtatagata	cataatatct	atatactctg	atagatacat

65521 aatattctata tatctgtata gatataaat atctatatat ctgtatatag tatataatata  
65581 taattataata atttatatatt atattataatt aaattataatt ttaattataa accataatata  
65641 taattataata atttatatata tttatatatta taattatatca cctgataatt taaagaacca  
65701 tcagatcaata tggagactttt tcaactattac aaacacagca tgggaagaagc ctgccccac  
65761 gatccaatta cctcccactg ggtccctccc ataacacrtc agaatgttggg gagtgtacaat  
65821 tcaagggtgac atttggglgg ggatacagcc aaacatatat acagtgttatt tagttgtatt  
65881 ctttaataata tttaactata catatacata gtgaatatat atatatattat atacacacac  
65941 acacacaaac acacacaatt ttatctgaat caatccctca taagtctgga aaatagatcc  
66001 caacccaagg ccaggatagc aagcatgtgc ttataacaaat aaMRttctta aagaagaatg  
66061 ccccgcttag gtaaatgtta ccaattcttc taatacccca aaaaattcaga ttgtttgtt  
66121 aaaaatctag atttttaata ttggtaaaac tttaataaac atagataagc aattatagt  
66181 ttctatggat atagggtccc algggctata gcttaacata tgggtctca attccagata  
66241 ttctaYagtc ttacagagct ctcaagltta gacctgaatc atgtctccag ttctcttaa  
66301 attaggttag ttaagaYgt gatcccaag ctacccctag ttgtccaggt accaggtttcc  
66361 aatagggtgaa tgtagagttc tacggatgag ggagctgttt gattttatgca tactcatgtc  
66421 agtatattgga ttatgccaaa tctgaattca gtagtttggg atacattcgc aagaataaat  
66481 gaaactataa tgaagaattt gtcaattata aagcctcaaa agtggggcat tttaattcca  
66541 tggttgYgct ttatggattt ttcatctctg gaataaggta taacacatgt aagtttggg  
66601 attttttctt tcttttcaaa cataaaagta aagttgtgat tggttttagta atagcacaga  
66661 cacttaaaaa gttaggtctt tattattctt ttgtgagtt taaataggaa aattctcagt  
66721 ccaaaaaaat gggcatctc tcttctgcca agtacaacat gtttggtaga ttgatatatg  
66781 atgtctcaat acagagaagc aaacagctgc ctccctaaa tctgtgtgat cagaggtttg  
66841 acctaataca ccaactctgt tatttgccag atatacagca aatggccaaa tggacatttc  
66901 acacatacca tgaacccaag ctcaattgcaa ttatttacta tcaaatgacc attcatatgt  
66961 ttctcaaaag tccgttccca aaatatgttt ttcatataact agtctccaat tatatacatg  
67021 ttacttttcta gttatttttt ccatttttta ttgtgattt ttgtgtgtac atttagtggt  
67081 tatacatttt tggaggcacat gtgagtcttc gatcacagca tgcgaattgga aataatcgca  
67141 ttatggagac ttgtgtattc atccctcat gcaattacc ttgtgtcat aagaataaa  
67201 atactacatt ttgttttatt tttaaaagta caattaaagt actattgatt atgtcagtg  
67261 tggttgtgcta tcaaacagta tgccttattc gttctttcta attttttgta ccaataagca  
67321 accctcacttc cctctcaact tcccactact ttctccagcc tctgttaagt atcttctgc  
67381 tccctgtgtc tatgagttca attgttttga ttgttgatc cccacaataa gtgagaacat  
67441 gcaacgtttt tctttctgtg cctggattat ttcatttaac ataagtatct tcaagtccat  
67501 ccatgtgtct agttacttta ataaacacaa acacacacaa tttttaaagt atggaaaaca  
67561 tcagtacaaa taactctccc tggccttcat aaatacactca aacactgatg aaagccagat  
67621 agcttttttt catataattc aaaaataagaa gtaaaagaca tttttaaagt aaactaaata  
67681 tatgcatatt tttttcaaaa tttaacttgt cagtctcttc tcaataaacta gatgttaaat  
67741 ttctataatt tttttcttta tgatagcatc ctctaaaata ttttgccat cgcagaagtt  
67801 caacaaagatt ttgtgaattg atcagtcatt aggtctaaaca ggataaataa ctctgactta  
67861 aagccaattt aagccctaag gcaatatggg ttgtgttca ttaacataatg ttgtcaaaa  
67921 aatgcatctca aactctttct aaaaaatttt cttaattgac atatccagag ttggccacctc  
67981 cttttgcact gqaattttat ccaatgtgtc ttctggaag ggacaaatgt tttacttca  
68041 catatctaca agttttcaaa tagctataga ttctcaagac aaggtctgaa agagaataat  
68101 aatcagttta aaaaatttca aaaaattagaa gataaaaata aagctctgaa cttgaaagag  
68161 atcaaaagta agaaaggtaa attgcttata gtgaaacaa taagaagaag ggctgttaaa  
68221 aagtttgggt cagagggcgg gtgcgagtc tgactcagtc aatctcagca cttgtgaggg  
68281 ccatgtgtgg ccgatacact gaggtcaaga gtctcaagcc aggcgtgcaa accccactat  
68341 ctactaaaaa tacaataaatt agcaggggtg ggtgcagatg gtctgttaatc ctactactt  
68401 gggaagctga ggcaggagaa tgcgttgagc ccaggaggca aaggttacag ttgagccaag  
68461 tctgtccatt gcactccagc ctgggtgaca aagcaagact ctgtctcaaa aaaaaaata  
68521 aaaaattgtt ttccacagag gtcaaccatgc aagaagattg ggagaggtc aggtgtgaa  
68581 ggaataaacac atcagctagg acagcacgtt aagtaaggtt tccacatatg ttgttaagctc  
68641 atcagagatg gtggcagcat tgaacttaga gatgataatt gtaacacatg tgcacctcc  
68701 aataatgaag ggaataatcc aataacttga tagatggcat ctaagtgaag ttctgaagc  
68761 tgaagaagat gtgtagtacg ctgcagactc cagtctcagg gctcaatgat gttctttaa  
68821 cagtgatgtt aaggctggaa gatctaaaga gtgtctcaag atatagatga taactataat  
68881 gagttaactc cagagaggtc aagaatttaga aggtggagag agagaacta gagatgggg  
68941 cagggaattat atgcagtaaa gaggaagaag gtgaggggtc gatggaagac aagatggaa  
69001 accttaagcc tcagacaagt gaagtgaatt attttatcc atgaaggtta ttcagacaag  
69061 aactcataca aaaaatgcacg ttcttgcatt tcttagcccc accctgtttt gcccataat  
69121 atagcaaaat tgggtgtgga aaagcagtag tctcagaaa ttaagctaat aattgtttt  
69181 ttatgggtgt gtatttttaa tgtgagatta aagtaaaaat caaaatatat ttagtatta  
69241 agattatgta actagaaaata ttctctctc tcttataatc taaactctta acaaattta  
69301 gcatctgact acagcctgaa taaatgccaa tggaaacata acaaattaac aaatagatac

69361	tgggaaatgt	caacattccc	agatgctgat	ggtcagacta	atgacgccaa	ttagaataat
69421	tgtctctgaa	ataaaacag	atgtcctaatt	acagagtggg	agcaccaatt	ttaagaatatt
69481	ggagcctcag	gagtagagga	agggaattgt	acagttaaat	aaagtgtttg	tgtactcatt
69541	gcgcctcaca	taattcagtg	tagataaaat	gcgatgtgct	aagaatttag	tgtgaaatgt
69601	acagaaaaat	ttaacaggag	agaggagaag	agactcccaa	ctctgcactg	aggagacaga
69661	agaagagttg	attcaacaga	tatttattgc	acacctatta	taaccagtaga	ccatgtctgg
69721	cacttagtgaa	cagtgagcaa	tagattcaaa	acttctcgcc	tttgtgagct	gacattctat
69781	aggaggagata	cagagaaaa	ataataatg	tcacatggt	aaagctaata	tcagaataat
69841	aaaattagag	taaggaaaat	agaataatga	ttaaacaggg	tgtgtaggaa	atgtctctctg
69901	ctaaggtgtg	acttaagcag	aatctgacat	ttgaggtgac	atgatcattg	tgtcttaaaat
69961	cgtgtattaa	gagactgcga	gaagacaaag	aggggaaatt	ccagacagag	aaagtataaaa
70021	atggaaaaac	ccagggaact	aaaaaaaaa	attgggggaa	ataagtgtgt	tttgaccccc
70081	taccattctat	tcctcctctt	aatggttgga	ttagatgagt	agacaaaccc	tacttctctt
70141	ctcttagagtg	ccagctagc	tcgtgagaatg	aaataaatat	agcatatgatt	aacaggaggga
70201	aagcgtactc	taacaaatgt	acatggcatg	agagccctca	tagagaaatg	aagactcgaa
70261	ggagcagctg	gagtcagtta	cttataact	gaattggaca	aagacgagta	aaacttgaaa
70321	tgtgactaca	ttatatatgg	aggcttaaaa	gataagagtt	attctaatga	aaattgtaca
70381	gaattctctt	agtcctaact	tctgttctct	gaagataaag	atgtcaacct	tcttctgact
70441	aaagagagct	tttcatggga	atttccacct	tttaagaaaa	agcccaaaag	aaactagtgt
70501	ctgtcactgt	ctgtctttca	agtgtcttca	acttaaatg	tcaattatgc	ccaaaatagc
70561	attatttttgc	ccactcggta	ccatgaaatt	ccatttagaaa	actgtctcct	ctctctcagt
70621	ccagtgtggt	tagaagagta	aaacacgtac	ctaggcctga	ctaattaaat	caacaaatgt
70681	ctgtacaagt	atgagtaact	caacatttgt	caaaccttgt	aaatccctgg	gcttgtatga
70741	aggttctcag	aaatctctac	tgggactcgc	agtgttttcc	atgatataag	tcttgagctt
70801	ctggaagtct	tcactctaga	accagggtat	gaaaaataag	caccacaaca	tagagctaaa
70861	atacaaaaag	aagtaagtct	gatcttcaat	tttgtttttt	ttttttttat	tattttctga
70921	gacagggtct	ccctcctctg	tgttccaggc	tggagtgtct	tagtgcattc	tcactcaact
70981	ccaaactcca	ccctcccaat	caataaatct	tcctctggac	gctctgggac	tgtctgggac
71041	tacaggcagc	ccactatgtc	cagctaattg	ttgtattttt	agtagagaca	gggttttacc
71101	atttctgcca	gctgtgtctc	gaactcttgg	cttcaagtga	tcacagtcat	cccttggctt
71161	tcaccaagtgt	ctgggattac	cagcgtgagc	cactgcaccc	ggcctgtatc	tattttctga
71221	aaactcagat	aaaaccaYgt	ttgcaggtaa	acttggccag	ttttttttat	acttctgggc
71281	caRtaaaact	ctctctttaa	catttagactg	gatattcctat	tattttggcaa	taagattctt
71341	gatagaaatga	ataatattgt	ttgtgtctta	tcagttttgg	tttctcaaca	caaatattgt
71401	agatgtgtg	gggagggggg	gtgctcaaat	aaataacatt	tattgtttag	tcagatgctt
71461	ggtagagggc	ctcttctctg	ctgtgtcttc	aYgtggcctt	tccttggcgc	atgcacagag
71521	agagaaacac	agagagagat	gctgtgtctc	ctctcccaat	tattgcaagg	gcactgtcaa
71581	tatgggagtc	cttgcataag	gactccttct	tcataactta	attgagctcc	aaaggcccta
71641	tcctcaaaata	tcactcacat	ggggattaga	gtttcaactg	attaatttgc	gggaacacag
71701	gcattctcata	gcacagttca	agaaaagaag	tgtcaaatgt	tatttttagaa	atttttagat
71761	cacttgggtg	agtatccccc	aaagacctct	aggaaaactt	tatttctgcc	attttaaagat
71821	agtgagtgat	tagaagcaat	gagaagaatg	gggaggtagt	ttcagaatatt	agagtaaaat
71881	gttgtagactg	gtttagaatt	cccacacaca	gaaccttgag	aaatcacctt	ctctggattg
71941	cttactcata	aaaatgtgca	catgtgttaa	gcacagtagg	tactgttggg	agcagcaatc
72001	tgtgggtgca	aaaatagatt	taacatttat	acattgttaa	agagaaacag	agcctacaag
72061	tagttaaact	gataaaaaca	gattagcggg	aaagagatct	catctatgaa	ctgggctcaa
72121	ttttgaatac	gaaaaggaaa	aatagggatt	tatagccaa	gagcaggtgt	gggtatgggt
72181	aggtggaaag	ttactaaaag	gtgacatcac	aggttaaagg	aatctcttgt	aaattgaact
72241	aacagagatt	ttgtcaaatg	tagggccagg	tgaYtgata	tcaagggtga	gaatagagga
72301	atttcatcag	atctgaaaag	tttgggattc	cagctccacc	tgggctatgg	agggccacga
72361	aggtaggagg	ccagcgaag	atggaggcca	aggttgagat	ctagtgaaga	agggagttca
72421	aggggtctga	ctaaagtgtg	ggcaaaagaa	gaatctttgt	cagttgtgac	acttggttgt
72481	atgtgtttac	aatcacttct	tgtttagact	gtattgccat	gtttaaataa	catgagagct
72541	taggggacat	tgtctcaggg	aaaaatgact	gtgcaactca	gaatcttgaa	acttggttgt
72601	tttgaaactca	actaaaccca	cactttttct	ttttctctat	gatttactat	tatcatcacc
72661	aaactactaac	ctatgtaaaa	ttataaatat	ttataacatt	gcaccataatg	aaacttttta
72721	acaaagtctac	ttaaatgggt	ctcaggtatt	gctgtactta	aaagtatgct	aaaaaacacc
72781	tcactcagaat	acacataggag	ctgtgttaaa	agtcagattt	ctggacattt	taaaaaacat
72841	ggatcaaaat	ttctgatagt	ggtgtccag	tatagccacc	actctagaat	attttaataa
72901	aaattccaagt	ctagtcttat	gcacactgaa	gtttgagaac	tatttcaact	tggagcagag
72961	taaaattatgg	ttagattttg	acaaaacaa	gagacttgga	tttatttgata	cccatatg
73021	cttttctacgt	ctttatacac	ctagtctcat	tttaattaca	aaacagacga	cttgagataat
73081	attttccctca	ttttacagat	gctgagatta	aatgagaaat	ggttatacat	tgcactcaaa
73141	tcacacatc	tgggcactgc	caaagctgat	aggaattcca	ggtctatctg	atgcttatgt

73201	aatagtccaa	tgggttcacc	tggcccgctg	cctagacaga	gctaatttat	caagtcaagg
73261	gaattgtcaat	ggagaaagag	ttaattcacac	agagccgtct	gcgcaggaga	caagattttta
73321	ttattgtact	atccagctctc	cataagcatt	caggagtcag	agtttttttaa	gaataatttgg
73381	caggtagagc	cttgggaagt	gggaagtgct	gattgggtcag	gttgagagatg	gaataatagg
73441	gtgtcaagtg	aggtttctctt	gctatctctc	gttcccgggt	gcgtacggag	aaatgctga
73501	gcagataacg	ggtctggggg	gcgtcaagctg	attcatccag	tgcagggtctc	gcagctgtc
73561	tcaagcactg	atcttaaaatt	tacaataagt	gatgttatcc	ccagagacaa	ttttggggag
73621	tccagctctc	tgagacacga	ggctgcatga	ccctcaaac	gtaatttcta	atcttctagc
73681	taatttgtta	gtcttccaaa	ggccaactgg	tccccaggca	agaaacggat	atcttccaga
73741	agggtcact	ctcagttttg	tWtagagtc	aaacattgaa	ctgattctct	ccccaaagt
73801	aggttggctc	acaccacagg	atgaaaaagg	acagcttaaa	gggttagaagc	aaagtggagt
73861	aggttaggtc	ttatctcttt	cactgtcata	atttccgtg	ttataatttt	gcgaaagggt
73921	ggccaggctc	ccttctgctc	ctcatttctt	ggatattcca	tgaaaatttt	ccatttccaa
73981	aagctattgc	taaaatatta	acagtaacta	tttgatttg	aacccaccata	atgcacacag
74041	ctaggctgat	ccaggaaaaa	gttagatgtg	tttcccgat	tatcattttt	ccaaaaaggt
74101	gatttgattt	atacctatct	ctttccctgt	gatttcagg	attctcaaaa	cttccacag
74161	ccaaagttag	tgatcactcc	taccctctgc	ttataattct	aaaaaggctK	attaaggtaa
74221	agaaatatta	atttcccttaa	catttaattt	tatccacatc	cttttaccac	ttaggaanaag
74281	tatggtggct	tgactgattt	aatgtgttgc	attaaacac	tatttcaagt	ataataagtc
74341	ttttatttat	gtactaacat	tctgggcaca	gaatcacaga	caagagactt	ttaggaaccc
74401	ttactattcta	aatggtctga	gagatattaa	aaaaataatt	acagatcaaat	ggattgtcatg
74461	tgcatactta	caaaagagata	agtaaaaaag	aatggggaat	tgagtgcact	aaatgtccta
74521	gaggacagac	atttatgccta	agagatcacc	aagcacagaa	ataatggaga	tcatttccaa
74581	aaaaaagata	gcattattcca	gagacctcga	gttgtatagc	acttgcctcc	gaattgtatt
74641	tctcagatct	tttccatctt	ctttaccocag	ttcatcaggt	accocagata	gtgacttccac
74701	ctgctgtctc	tttcagagac	caaaaggcag	gcttgttact	tacgtctcag	tttgaaccc
74761	ctggcatgag	ttattcagat	tttaaaaaag	ctatttctac	aaaaggcagt	ggatttcata
74821	tacttctctc	attgtcttca	attgcaccaa	aacattttga	attagaattct	aaaaagaggt
74881	tctaagtggc	cactggaaaa	atacttggata	aatattttgag	aattttctcc	tactataatt
74941	ttgagaaaaa	attgtaaaaa	tggtgttgaa	cttaaaaaac	agagcgtgaa	tagctatgat
75001	ctcttgagag	ttttatttat	catacagggg	ttatgagctt	tatgtcaaac	tgctctttgt
75061	aaaaaagata	tttagataag	gactcaaaaca	ttttcagtta	ttttgttgca	cagcttttaa
75121	ttgtgtcagc	atgcttttgt	tactatgagc	atttaccaat	tttatatgaa	ttaatcacat
75181	tctcctctgc	catggatact	cctgttctct	taccocacaa	tttaattact	ctgtcttctt
75241	taagacaata	ttttctgaaa	acaataccgg	tctcattatt	taaWgtgaa	acattttctta
75301	gtttttagcac	tctcatatgt	tgattttgtca	ttttcatgac	ttataaaatg	ttagactatt
75361	actctataac	catgggaagc	tatttcaagt	gaggccagaa	aatgggaata	catattctca
75421	aaaagaaaaa	attcaacaa	atgttccatt	catctttgct	ttggacaaaa	catgtgtata
75481	ccatatgaca	aattaaagaa	cactaaatta	cagctttgaa	aatacatagt	tatctataaa
75541	atctctgact	taagtccaca	ttaacttttc	ccatagcaca	gatttttaag	tagccttaat
75601	gatatgaagt	gaggcagaa	ttaggcattt	aggtgcttta	ttaagaaagc	taagattcag
75661	gggaataatt	ttaaacaact	tggaaatcca	tcagtataaa	tatatataat	tttgaatcaa
75721	aaaaatatatt	taaaaaata	ttgaatgcaa	cagtagtggt	gcataattct	taattgacat
75781	atttttgttt	catctaaatt	ttctcctttt	gacatctcaa	aaagatgttg	cactctata
75841	aaaaaatcta	ttcaaaagcct	tatctcact	gtttttcaac	ctggttagcta	ctgtgagcta
75901	ccacttttact	ccatttcaatt	aattcatttc	tttaactcatg	tcttagaact	accacactag
75961	ggctactgct	atggtttgaa	tatgtctctc	aatgttcatg	tggtgaaact	taattaccat
76021	tttaagtgtt	ttaagaaatg	ggggccttga	ataggtgatt	aggtttatag	ggctcttctc
76081	taagtggatt	aatgccatac	ttgcaggagt	gagtttctta	taaaaaggat	gatttttgtt
76141	ttatttcact	ctgtcttgca	tactcacttc	tgcttctgc	ctctctctca	cttctatgat
76201	tttgcaggat	gccaatgtca	ttcccttgga	cttccacgct	ttcagaactg	tgaaaaaatg
76261	atttctcttt	taataaaatt	actctacctg	tggtattcca	ttatagtgca	gaaaaatgaa
76321	tgtgtggttg	tgtaatatat	taaaataata	caccacacat	tgctattctc	aaacccattag
76381	ttctcatctt	tcacttttct	agggccctta	tttccacaa	tatcctgtag	aaatgttagc
76441	ttccttgact	ccacatttgt	ttccatatca	atgagctcct	ttgtgactct	acttttgtct
76501	tgatgcatca	tttcaacaac	tttctgccaa	tactctgaac	atgctcatct	atttctctct
76561	ctgatgcaaa	agccctgaaa	atcttaagca	tttttcaaca	tacaaagcct	acttttagat
76621	tgtattttcc	agaaaagaat	tggtgcataa	gttagagatt	atctgggttt	gggtctttac
76681	aggttctctt	caataaaaaa	ctgtctcttc	aggttagtcaa	atataatgat	gagtaataat
76741	ttgtttagtgc	caacaagggt	cataattcaa	atatttaaca	acagaccaaa	tgaaataggc
76801	accataataa	aaatgtcacc	tgtaaaaaag	agccaatatc	agctttttat	gaagacaaatg
76861	agaaataaact	ccagaagaag	gtggggagac	ggaaacaata	aaataaaatt	agaggagata
76921	ccaggtgagc	taacagccaa	gaaaaatgca	ctgaattgca	agccttttga	atgacaacaga
76981	acaaagtcag	atctaggaat	attagcagca	aaagactgtc	agccttctct	ctgggtgctg



77041	ccataaaatg	acatagctat	aaccactttc	tgcctttgta	tattttctct	caagatcaaa
77101	atcttttagt	tgttgaataa	tgtgtctccc	aggtctaggc	agctctccc	accctaaatt
77161	catgaaccaa	tcactatggg	catgtgagat	ggttccctcc	actggtctag	ccaaagtctc
77221	agggccgccca	agccatcaga	aaggtgaggg	aattctctga	agattgaaaa	gggataagtg
77281	gatgaagaca	agaggcacta	cagcatcttt	tgtctgcaat	gaaaatttct	cttttgatgt
77341	gcttttMagg	agccctctcc	tagtatattt	agagagaaat	ctttgtctca	gtttagagttc
77401	aagttaaatc	acagaccaaa	aagctattat	aaactctctc	tcaattactt	tccagatttga
77461	ccactttcca	ctggctattt	actctctgtg	ctgtctctct	actcagcttg	cgctgttaaat
77521	ctgtgtactt	ccagcagctt	tgacctttct	ctttttttgt	ttttactaca	gcaaaccttt
77581	ccctacaayc	tcatccacac	ctctggcttc	agcttcagat	tctattctga	ctctccaac
77641	tgtgagacag	attaatcatt	tgcgttttag	actcttaaat	gcaactgalc	tattgaacat
77701	tccctctcca	gataatgttg	gggaattttc	cacagaaaat	tcaaaattta	tgtttccaaa
77761	taaaattcac	catcctcttc	ctctctccat	tcaattctct	tctctctctg	gtcatcttag
77821	ctctcattcat	gcgacacca	tctgctgaaa	aatccaatct	gagagcttgg	aattcttcta
77881	gactcaaccct	ctctctcaaa	ctcctttgtg	caatccacca	gcaaacgatg	cgagttctac
77941	tcactgtctat	tgcctcagta	aaagcactcg	tcaattctcc	taacatactt	tctctactcc
78001	tgcgttcaatt	accatcaata	tgttttcaaa	actctgctag	aaaaaaacta	aaagggcaat
78061	tgttctatgt	cccttaaaag	gctgcctaca	aaataRcact	caaaactctt	gttctggaaa
78121	ccaagatctat	ctgaagtttg	tatatctccc	ctctatctag	tcccagctct	tgcactccct
78181	ataaacacca	cttgaagtgc	cagctgcacc	aaacagttct	gtttgtgaaa	cagagctgtg
78241	catgtgtctc	ccctagacaaa	gacaaatcat	tctcactgag	gcttgtccaa	atatctgtct
78301	tgtacataca	aatatattgt	gttttaccat	acagtgggta	ttagtcagga	tctctagag
78361	aaacagatct	ctagagaaaa	taagatacat	acatacatat	gagacaaaag	gaagagaaaa
78421	tgtgagagaga	gagagactta	tgtgaaagta	agatactggt	tcacacaatt	attggaattct
78481	agaaatccca	caatctgcc	tctgcaagct	ggaaactcag	gaaacccagt	gagagccaaa
78541	ggggtatatt	ccagcctggg	tctgatggcc	tgagaaactaa	ctcagctgtg	gttcagaaat
78601	gatcaRtaat	ccagctaaag	caactcaggca	gtgagcaaat	tcaacattct	ccactctttt
78661	tctctataca	caactctcact	ggatgatgcc	ctctcacaat	gagcaaaagc	actctgttta
78721	ctcagttctg	taattcaaat	gctagtgtct	tctccacaca	cccttcataag	cacactgaga
78781	aaataattgt	aaccagatat	ctgagtatcc	catggcccaa	tcaagttgac	atacaaaatt
78841	aagcatcaca	ctgagtttga	gKatatata	tacacaacag	tataatacgt	gagcaggtgt
78901	ttttataatt	tacagatttct	ccatgggtgaa	aatctccagc	attacctgtg	gaatagattt
78961	ttactctttt	acagaattgt	actattttct	ctttctggat	tattatgtgt	agtggtagaa
79021	aatataatgaa	atttgagaac	agacttaagt	tcaaaaaata	tcatttccct	tattatcttt
79081	tatgaacttt	gataatttcc	tcttagtaaa	atatggatga	tatctacttt	gtacRtagct
79141	gtcgcRctta	attcagctaa	tagggataat	gctcctaaca	tctcagattg	gtcagaagt
79201	tataaacacg	ggactattct	tctccatgat	gattagcaat	ggagtgtcta	aaaagaaata
79261	agagaaataaa	atatgagaca	taaaatatta	aagagtgtac	attgaaatgt	tggaaacctg
79321	agagaaatcaa	gcaaaaattt	attcacaaca	gtaagaatgt	ttaagtggta	gctagctaca
79381	aaatttaaat	ataagtcaac	aactgtcttg	tatttgaaga	ataattaatt	tgaaaatatt
79441	actcaaaaag	acctcYatga	gaactaagta	taaaatactc	aagaacactgt	ttaaaaagaa
79501	atgtgtataa	tatttgaaga	aaacaatgaa	agctctgcat	ataaaatttt	aaatgtatgt
79561	gccacgacgt	gagtaattgt	tttgatgaat	atcagttccc	tctctgcaRg	cttttagcat
79621	tacaggctag	ctcaaggca	acactacagc	ccctctgcac	agctacacta	agctctttct
79681	cataaagcag	atgaggcaga	tgtcatcaac	tgagagactg	gaaaaatgca	accaatttat
79741	tgttagggaa	atgtcaataa	ataaaaattt	aatgtggaaa	ccagtttaaaa	ggcccatgaa
79801	agtcattttc	taccaggcaa	gatcaaatcc	tgagatcaac	tgccctcaat	tgctagagtt
79861	cggtgatttt	aagagaaaata	tataaagctt	agctttcaac	aaatttaact	tctacatttt
79921	taaaagacgt	tgaggcttac	ttttctgaaa	atgtgggggt	taagaaaaaa	aaacaaccaa
79981	gcaaaatagaa	aatgttagtt	ctttagactt	acagaagaaa	agcatgacc	tgtttagaaa
80041	gtgttggggt	cccttttaatt	tttttcaag	gttttaggtt	ggttttRatga	catgatcttg
80101	atatttatatt	ggtataaaRc	tgcatttgc	aactatgttg	tattagataga	ttatttgggt
80161	ttttgaattc	aaagtctttg	agaaaaaat	ttaatcataa	gctaaattca	tcatatttat
80221	gcacttttta	ggaggagttg	tcaataagata	gatttgaatt	tacatttagt	tttagattgt
80281	aatgataatg	ataagcttcc	taaaagtggt	tggaaatttg	tctctctcat	agatcaatct
80341	ctttttcac	tcgagtga	atttcaatt	tgttcaaaag	cttttgaac	aggtctgtgt
80401	atttctcttg	caacgtatgt	tatctcttt	ttccctaaac	ccctaatctt	agctgccttc
80461	actcccaatg	gaaaaaagat	gttttccagc	aacctgtgaa	tctaatgttg	actatatttt
80521	aaggtcaact	tgatgcccta	caaataaatg	tcttttgttt	aaactttctg	aaagaggttt
80581	atagcccttt	tgtcccttct	tctctttctg	ctagctgaaa	aaaaattctt	tccacatgta
80641	cagagaataa	aactgttatt	tattttaagc	catgtgtatt	ttagtattact	gtatctcaag
80701	ataacttagt	tattcaagtt	gtcttagttt	actgtcaact	aagataacag	tctccaattac
80761	aaacaYattt	ataaaaataa	ttaaaaaaaa	caagaaggct	accttccagt	gagtaaaaa
80821	aaaaattacc	caggagctaa	tgtattttta	aatctgtctg	agttgtttct	tgatcctagt

80881 atttccgtgt tgggtttatta tctaaattgt ttgcatttat cagttactgc tgetatgact  
80941 caatcatatt aataacatga tcgaagagctt cctattttaa atctactcga atgcgatatt  
81001 gcctcactga taattattcaa gtctggccaaa caggataaga aaaaacaaatt ggaatagaaga  
81061 aactcataag ataagtctaa cactgcttaatt gaataattta gcagcttgtgtR aagaagaaggaa  
81121 agctcataag atactgttact ggttagaggtg gacaaataag ttaattacct aggaacactg  
81181 tcagccatag atacagtaac taaaattgtgt gtgttaagct tatataagaaa gggctgggaa  
81241 aaacaggata aaagtaatat tatgatagwt gaagtctctga aagggttgtta agcatgttat  
81301 cactcctctc agttgctcct catttgcata aactaagact caa tagctta ttgtgaactga  
81361 ccttttatgt actgtctctg agagaacaca gtaaaagact caaacacacta ggaagttaag  
81421 caactggaa tctctctatg catggcaggt tatagtctgt ggaatgcgtt ttgttaattac  
81481 tctttttgga gagcctcctt catgtctctc aatacacacc agtttccatt cctgtgttatt  
81541 acctttggcta tggccaggtga tatactggac actgtgtgtc acttccacct cctgtctcaa  
81601 agatcttact tggtaattct gtccacaaa cactccaaat ttcaacagct acatctgtgc  
81661 tctccctYac catggctccc tactacaaa tccaaacttc cagaagataa ttgtctgggg  
81721 aactgactgt tctgtgagct taaggaaaga ttgtagaatt aatagcttgc ttataataga  
81781 ggtggctgt accatcaaaa gaacaatgaa gtattatagc tctggattca cacttaga  
81841 aataacaaaa tggatttaact cttcagaaga aaataacagt gaacagctga taattcatca  
81901 ttctcatgt tgcataactt aaattgataa ggggtctaaaa caaactttat cttgtctgtg  
81961 gtagaagaga ttaacatctt agacatttaa gaattaaaaa taaaagctag ctgcttttaat  
82021 aatccactga cagaaacaaa agagagattt taagtgggct gaacactcag aagagtgtga  
82081 acaatgatta atttttaaaa acatacattt gtgttttttt ttcaactaga agaaatctgc  
82141 attttgtcag cctctcaatt ttttttaatt attattataa ttttaagtt ttatgggtac  
82201 ttgtgcacaa tgtgcaggtt agttacatat gtatacatgt gccatgctgg ggtgtgcacc  
82261 ccatcaactt acattttagc attaggtata tctcctaagt ccatccctcc cccctccccc  
82321 caccocaaaa cagtcocccag actgtgatgt tctctcttct gtgtccattgt gctctcattt  
82381 ttcaattccc acctgtgagt gagaacatgc ggtgctggcc ttgtgtctcc ttYgaattgt  
82441 tactgagaat gatgatttcc aaattcatcc atgtccctac aaagagacat aactcatcat  
82501 tttttttgct tctctagtat tctatgggtt atgtatgcca cctttgtgac aaacctgaSa  
82561 aaaaacaaga atgggggaaa gattccctat ttaataaatt gtgtcgggaa aactggctag  
82621 cctatgtatg aaagctgaaa ctggatccct tctctacacc ttatacaaaa aaattctcaa  
82681 gatggattta agacttaaac gttagacctta aaacctataa aactctagaa gaaaacctag  
82741 gcatttacat tgcagacata ggcactgggca aggacttcat gtctaaaaa ccaaacagca  
82801 tggcaacaaa agccaaaatt gacaaaatgg atctaattaa actaaaagac ttgtgcacag  
82861 caaaagaaac taccatcaga gtaatttttt taaattcaga aaattctcat tctgtaaagga  
82921 aaggggcgct accggcatcc ctacagtttt tgtcagaagc tctgatctcc ctgctcttt  
82981 attccctttc tccaaaactc tctccacact gatccagat aactctttaa actggcaact  
83041 tgatgtgttc actcctttgc ttaaaattct tcaagtgttc tttaacagat aaaaatctaa  
83101 tgcctttggc tgatgaaaat ccttttgcaa tgcacactgt tttaactcag ggccttctct  
83161 tctcttctct tttccctctg tctctgtgtt tccctctaga gtgcagctgg aagtcagctc  
83221 gctggtctct ccgtgggtga cgactccagc cagcacaaat tctcttgat ttcttccacc  
83281 cctcttcttt tccaggcRtt ttcaactcag cctgcccaag gaacagctgc cactctgtcc  
83341 aattctgtgtg tctgtcagg cctctgtctc tccacttct ccttctcttt cctgaactcc  
83401 gatcttttca gcatccctcc catttaaaac aaattattag agtcagatat ttctctctt  
83461 aacaagccca catttaaaac aaataaaca ataaaaatc acacacagct cctcaatctc  
83521 agagagcttg tcgattaaac aaataaaca ataaaaatc acacacagct cctcaatctc  
83581 catctcctgt ccttgggtga ggggtgatg caagacaata tctgctctgt tcttaagact  
83641 gacggaaaat tacacatata tatacatatt catatat tca tacacatata cctacatata  
83701 tactgttaca tgttatatg ttctctatga tatattgtgt ttatcattgt cctttaccag  
83761 atcatgttac agaggtaaaa actgttataa aaatgggaggt ttactcctgt cctttaccag  
83821 Rgagtgattt tagggcaaga tgcacacttt tctctctctc tcttctaaact cctttgtgct  
83881 ctctctctct ttgtggctct ctctcagctca ttaattttat tgcataaagt ttgttttaga  
83941 gtaaacaaaa gacatgcaaa attcttctct tcaataaaat ttgtgtttaa aaggagagga  
84001 cagacataac acacaaataa actctgaaaag ccagcaggtt aaagtgtatg gcagctaggc  
84061 agtgaccatt ggggaaggtaa ggggtgtgat ggtgtttgag gtacatttct caggcagctc  
84121 Yccaaattgt gattgaaagt ttctctYgat ggtggggagaa gccataatga tttgtttgtt  
84181 tccagtatagc attttctttt acatatattat tctataatgta acttttttca actttttgaa  
84241 gcagctttacc acttaagccc catactgtca tgtgaaagct tgaagcccat taagatagac  
84301 ctcaaacata agaacaaaaat aagagaccag acgctgggtgc tccactctgt aatcccaag  
84361 ctttggggag ccaagggcgg ttgatcatga ggtcagaggt tgcagaccag cctggccaaa  
84421 tgggtgaaacc cctctctctac taaaaataca aaaaatagcc aggtgcaggt cagggtgca  
84481 gtaatcccaag ctactcagga ggtcagggca ggagaatcac tgaacccgct gagggttga  
84541 tRagcccaag atcRtgccac cctgggcaac ggcactcagc cactctcaaa cactctcaag  
84601 aaaaaaaaaa aaaaaagaaa gaagaaggtg gggctttgaa ggaagatgct gactctgaag  
84661 cagcccaact agactaaag agtgatttga Rtgtaacaaa ggaatccctaa acagctgaag

84721	caaaggRaca	gatgaaaggg	tttcgaggat	cttcataattg	aaggaaacgca	agcttagccca
84781	acttagcatg	gtttttctta	gtaaaagcttt	tccataacca	agatttttccc	agatttttccc
84841	actcttagaca	aattcataaaa	ttgggttggtt	tgggggggaca	gttaaaaacac	agaataaata
84901	atgttcaaaa	atactctctt	tcatcaatct	tctgtaaaag	ctaaagagcat	ataacttaat
84961	tacaactgtg	tagaggccaat	acgatggcctt	gctttctcgga	attactgtttt	tcttcattat
85021	ataaactatg	ataaagaaaa	gatggataac	tttttaacttt	ctacttgaaa	ataaatttga
85081	acatgcagaa	agttataaga	atagttcaaa	gaaaaacaag	atctgttttta	ctcagatcca
85141	accalttgta	ataattttgtc	catctgcttt	atctttgttta	tttgtctaat	attctacatt
85201	ttttctgaac	tactaggaat	aaattacata	caacatgacc	ctttaactat	caattgcttaa
85261	gtgtgtatct	ctcaagaatt	agaaatattct	cttgcataac	tacattgttt	gactacatttta
85321	agaattctact	aacactattgt	tcattttatc	aactgaacaa	agacctgtgt	agctttttgct
85381	ttgtttttgt	ttgtttttgt	tttcagtaact	ttcagtatag	gatccactct	aggataccac
85441	agtcataatta	atttttcaggc	ttcttgaatc	tttaactctg	aagagtttct	cagcttttacc
85501	tttttataaa	attttttgaag	aacatttctc	tccagttaat	aaaaaggatca	tcacttggag
85561	ttttcttaatt	gtttctctcag	aactagattc	atgttaacat	tctttagcag	catattgatt
85621	tgcatactct	tagcaagagt	gctacacaag	tcgttttatg	gtcttcttag	gacataYat
85681	gttgagacat	atcatgttca	tctattttcc	gtgttaatat	aaattttgat	acatgacca
85741	aggtcacagc	agctactctc	agtggtgtagt	tattatttta	tctctctgca	tactataagc
85801	aatccatattg	gagatacttc	aagactgtgc	aatatcctgc	tctcataaaa	aaattttccc
85861	tgtgtattat	catcatctcg	tgattcttgc	atgaactgat	ctctattgtg	atggttgcgc
85921	aatgattgat	ttcatactcc	agcatgccct	gcacattcat	cagtcattcac	tcactctcta
85981	caatgcacaa	gagccctccc	ttctcccag	ttatttactc	acatatttta	atacagcatc
86041	tacttatgga	ttcctatttg	tgcaaaagtt	ctgaatttat	tactatttat	acttattttg
86101	attctcaaac	ttaccaatgc	ttggcctgtg	agctcctgag	tctttgtgac	tattcccgaa
86161	tgtgtatata	tccttaattg	aatgcattgt	aggaccatga	aaaccaatgc	ttctcaacgt
86221	ctggctctct	ttgtcctgc	atatatcacc	cagggtgctt	gaaacaatga	tataacgtct
86281	tcgtactcaa	acctactgga	ccagactttt	tgggggatggg	caaacctaga	tgatttttat
86341	gacacataaa	aatcaaaagtc	gatgacatag	ataaggacag	agaaattcta	taccagatag
86401	agttaccagg	caaaagaacac	ttttccaaga	cgaatgcagt	aggggtcaaa	acttatgtta
86461	caaaagatag	agattaaact	cagctctgct	gaaacaaaag	gaggcaggaa	ttttaaatgc
86521	taggggtgagc	ttatggaaaa	tactctgaag	atttttaaga	ggaagtgtgt	ccaagtatg
86581	asgttggttt	tattctctaa	ttggcactta	tctcatttag	gttctcatct	taccatgag
86641	actcagagat	gggcccattt	tttcttgatt	acgcttaaaa	ggaaggtgcc	cagatcttYR
86701	agaaaagacat	tactgggttg	gaaaagattt	atattctcaa	gggcaacaa	ataactttac
86761	aatcagaagt	tttctagtaa	cttctctaaa	aRagggaact	tagggggccca	cagtcaggaa
86821	gaattctgtc	tRaagtttag	ccaaaactggg	tggaatggtt	agggctcttt	ggccactctc
86881	cacaaaagca	gagactaaga	ctattcttgt	catctccgga	ttccctgtag	ctagctcagt
86941	ttctaccctg	ttgtaagctc	tcattaaata	catatgaaat	gaatgaatag	athtaattgtt
87001	cattttatga	tcagagtcct	cagggaattt	gtatttttat	tgcaRaattt	tacttaaaat
87061	ctgtagggtg	ccctctgttt	ctttaggcaa	tatttctcaa	ccgtgttcaa	ctctaaaaat
87121	ctctctggag	gcttctaaR	aatcaggttc	cagagctgta	ccctaagttt	cactaaatta
87181	aaatgtcagg	aagttagagag	gtaagtattct	gtatttttaat	gcaattttaa	aactatacca
87241	ggggagatga	tttaagcatc	tatttctcta	acatctcagc	accataggat	catattttta
87301	agagctcaac	ctgtatttat	ttcaataaga	tgtaagggtga	ctctattcag	ctcaaaattt
87361	agaaMaggat	tcagttgcca	ttttccaagtc	tggaggtgcc	tggtgtgtta	tgtgagcagt
87421	ttgttaatca	taccctcttc	tgtgttagcc	tcgggtttta	tgggaagcat	aaagatcaag
87481	ctcgtctgaga	ttttaaagtc	agcctccgct	actcgtgtgt	tctgaaacct	agagcggtgt
87541	agctagttagc	atttgtatca	attactatta	aagtaaaaagt	ttggcacaac	actctctctt
87601	gagtgaagctg	gtaatgaaaa	aaaaaatcac	aaaaatgcttt	tctctgaaca	catcttagat
87661	ctctacccca	ggatgggtgg	caaaaattca	tctctgagag	gctatggcac	atgtcccttg
87721	ttgttaccoc	ttgaaaacta	gttgagaaaa	ttgctctata	aatattttca	actgggatta
87781	aaaaaagaaa	atacccgcta	atcagatcca	ataataaacc	tcttacagaa	tgcctctaaag
87841	aaactctaaa	tcttgcatca	ctgtgaaat	taaaagttta	ttccaaaagct	gtagggatta
87901	tacttaagaa	aatttccact	gctgtgttaag	agcttactaa	tttccaaagca	tgcctctaaag
87961	tagtccctaat	gcatactctt	atttaataat	cataatggcc	acataagtta	gataatttat
88021	atgctcggttt	tacagataaa	gacactgagt	ctcaaaagagc	ttaatttagt	cagagcgttc
88081	catattgtcta	gtattacatt	agtgcaaaag	taactgtgtg	attttgccatt	taaaaaaaa
88141	gtaatggcga	aaaccacaat	tacttttgca	ccaatgtaat	acatgcggaa	cagagcgttc
88201	aaatgcttagg	agctgtcttc	agagttagca	ccaactgaat	attttttaat	gggtgttctc
88261	aactcagagac	aaattttgtcc	cagaagacat	tagcagaact	tgcagaactt	ttgtttgtca
88321	aaggagagtg	tagaggggtg	gYgtgctatt	ggcatttgtg	gggtagaggg	catggatgct
88381	ctgaaacacg	ctgtagtaca	cagaaagcgc	tccacaacaa	gtaattttcc	agcccaactt
88441	atcaattagta	ccaacactta	aaaaacctgg	cacattggaa	ataaaagcct	tttgaggNta
88501	ttttatggaag	tatacattat	atatacatat	atgtatatat	tatatagagt	ctctcgagaa

88561 ttgttcacgc cgttaaagag gaaaagctRt actgtgcgatt tcaagctcta cgtagtaata  
88621 gaaaccttga gagaacactg aaaaactaaa ttcctgtgga acattaaaaa tctctlaattt  
88681 agtttaagat gatacctaac agctttccag aaccgttgat tagaaaagtg tctgttttcca  
88741 cagagaactc ctgagtcacac tctaccagat gtttttattc gtttttcccc caaagatattt  
88801 ttattttcag agcaagactc attgaaatgt taatttcggt aaaaattgagc taacttagtg  
88861 aacatctctt cagtggtata ttcctttaga gctaaagtg tatataaatt caatlaagat  
88921 gactccctgt atttgaattt aaataggttaa aaglaagcaa ttcgcaatagg gcaaaattctt  
88981 aagatgttat taaaatcaac ataatttttt atttccctgc aataaagattt tttgaagctga  
89041 tgaattgaaa tatatttctaa ttggaaaagc aataaagagt aatgagagctc tgcctgaaat  
89101 taatttttaa tcttcttttt aatgagtgca caaccaactc aataaagattt gctcaagaaaa  
89161 tattttatgc ttacttaaat caatgagacta aatgaatact tcttttctca ttgttatcttt  
89221 aaagtcaata taccataagcc acaagcattt aatatatat taaaaagattt gattctcttc  
89281 taacacatca tctgttagagt ctttggaaaa aatgacgata aataaattt cttctgtgat  
89341 ctaaatcttt ttgtgtgtaa cccaatattt ttaaaataa tcatcgtagt tatcatgtgc  
89401 taacaaatatt cctgtcaaat gtaaaatgct gcatgtaaaa tattttttcaa attttataag  
89461 gttatgcagt agcaatcaat gtgcactttt atttgcctca agtaagaaggt tattaccaag  
89521 caaagcaatt tttagagcaag aagagcttta aagtttggc cagtaacattt tgttctcttt  
89581 agagctgaga acactgatgt cagagagagt gacttttcca gacttagata gtatatgggt  
89641 aggactggct gttaagccac tgattccatg cagctctctc tactactattt tgcactattt  
89701 cccaagaagt aaaaattagt tttttccctt ttcaaatctt agaatgccaa tatatgtgt  
89761 acgtaccaaa ctgactgttta tggtcagcag atgccccaga gctttaaactt ttttgggaaa  
89821 ggaataaatt aaaaagataa gaaatggtaa tacaagtcac aaactctggt agagaagaaa  
89881 catcttccag tgggatgaaa acaatcatct ttaaggggaa tggttttatct gactgactct  
89941 tccaagtata aatgggattt caagagaaaag agtaaaaagga caggagctttt cttgagtggtg  
90001 agattctatgt gaacaaatat gcagacgtag gaaattatag gcatagcttg gaaaaaacga  
90061 aaagacttgt ttgtctgggg tatgtgattt gcataaagaa tcatagataa taatgaagaa  
90121 agattaagtgt gaacctgtgt actctgtgtc cttgcaatct cctgtgtgatt gttagagatt  
90181 ctaattgcag ccgggtgtga gtggctcaca tctgtaatcc ttgacttaac ttgagcccaag  
90241 cgggtgtgat cactgtaggt caggagtctg agatcagctc gggcaacatt ctgaaaacccc  
90301 gctcttacta aaaaacaaa aatcagtcgg gtgtgtgtgc acactgctta attccagctta  
90361 cttagagagg tgaggcaggg gtatcctctg gacccggagg gtggagggtg caggtagctg  
90421 agatttaccac actgcactcc agcctgggca acagagtcag cctccgctctt caaaaaaaa  
90481 aaaaagaagag agtctctcatt ccagctagaa gtgaatgatg agcagaattc acaagaatgg  
90541 ctgcttgaat gtgacccctgt caaaggctat ttgtcacta gagcctaagt gaaataatata  
90601 taatttgaag gaaatttttt ttgtgttttc taggattatt ttattttttg ttgtcagcaa  
90661 ttatgcgaac caggttaacat gatgaaactt tgggaagcta ctaaaactgc tgcatactga  
90721 ttgttgcact gaagttttta aagtttctga tgattctgta taccaaattc attgtatttc  
90781 tgacattttgt aggtcagatt ggaactacaa aagtaatga attatggagt tagttctgct  
90841 taagtttgtt aaaaattttt gagtttacc gtcgaatttt ttgtgataga gaaaaatga  
90901 caattaaagt gctagaaca taataaatta attaaaaaaa gaatttttaa taaaaaattc  
90961 aaagaattctt ttgtccattt ttgactttt tattctattt ttattagaaa cccagggata  
91021 ctattactgt gaagtttagg tgctctctgg gaaagatttt aagattttctt tctcttacta  
91081 gaaatttctt tttaaagag atgctggcaa ggagaatcat gtgggtcagt acacttttca  
91141 ccgacttcca ggaatgctga ggcaagagta caaacactga aataaaaagt agatttttaa  
91201 agtttgaagg aatatattca attggaacct ttgactctgg gaatttgaat atgacaaaa  
91261 gttccaaaat taatgatagc taatgtttat ttgactctgg gaatttgaat atgacaaaa  
91321 aggcttatat gtaactaagc aataacccaa tcaggtattt tatagagagg gaaacgggag  
91381 cacagcgaga ttgagtaccc tgccaatggt gaagctcgaa aataaattga attagactga  
91441 ctccagaaaa cattctctta actattgtgc catgtgttt cttagagcat ggggagatag  
91501 cagctccag gaaaaagtg aagaagcgtg gaaagagcta ttagattaat taactctctt  
91561 atcaagaataa ggtcaaaaga accctattct ttgagaagaac agaatagatg taaaattcat  
91621 aacattagaa tattttttta taattctctc tgagaagaac cagaactcag taagagttagc  
91681 tgactctact tgaggcagaa aatatcgagg tcagctctga cattttgtta ctgcaagaaa  
91741 acatgtgatac tttaaagggt gctctgtggt gaaactaaca atttagacc attggcctga  
91801 cgttgtctgt actgtccaaa cttagacccc gggttattgt ttgagagacc gctgcgcaac  
91861 ctttaattat tagaaccagt tgacagccag gcacgtgtgc gtgagactgt aatcccaaga  
91921 cttgtggagg ccgaggtggg tggatcacia aaaaatttagc caggtgtgat cctgcgctcc  
91981 atagtgaan cccactctga ctaaaagtga gggagaaatc ctggaacctg tgaagcagat  
92041 tgtagtccca gctactcagg aggcgtgaggt gggagactga cttgaacctg tgaagcagat  
92101 gctgcagaga gccagggcca tgccactgca gccaagcctg gtcagactga aatgactcat  
92161 ctgaaaaaac aaaaaacaa aaaaaaaac tgacatggag aaaaactatt aatgtgat  
92221 gtagctcaat atagcaaac aatatatagc atgtctcatg gactcattgt tcttccctaa  
92281 agttcatatg ttgaaacctg aactttcaaa ctgagtgcca acagggcctt tggagaataa  
92341 tttagtgtcat gagagtagag ccctaattat acatgggatt agcttctcta taagaagagg

92401	catgagactg	atgggtgtctt	tctccaccat	gtgaggataa	agcaagaata	tgctcaatctg
92461	caaccacaga	agaggagactt	caaccaattg	cyaaatagcg	ggcaacttat	cttgagactta
92521	cgaccctcca	gactctgttag	aaataaattt	ctggtgtgcta	agtcacatcat	tcocatgatgt
92581	tttctgtcta	gcagcttgaa	ctaagacagt	atcctagata	aaaactatcat	gtgagttatt
92641	ttatgggtac	aaaatgttcca	acaacacaaa	tatatgacca	atggtgtctac	agaaatgacc
92701	acacagcctc	cccatcaaga	ttcagggtgtg	aactacagtg	aaattactgtg	gtgaaactag
92761	tcccctgtatt	cagtatcaga	aacatccctc	ataagagaat	taaatccaag	caggagactg
92821	ctgcaggaca	ataagattttt	ctatggagat	atgaccttgt	tgctctcatgt	agaaagcaac
92881	atcttgtgata	tgtgtaaaata	taacttgatta	aaagtgtagc	ttgtttgctt	ttttcttaag
92941	ttctctctga	ctactgatatt	ataccctgat	gacaaaagcg	ctgggcactgt	caaaaactgga
93001	ttctctcaca	ccaaaagacc	ccttccctcg	tcaaacagca	tcacaagcac	tatccgcacc
93061	tattgatat	tgtgcaggcc	aagaagggcg	tgctgtgata	catgcaatctg	cttcttctcc
93121	agggggtcaa	cccataccoca	atactttttt	tgacttctgt	tttaattaaa	ttggaaatgt
93181	aatacaaat	gtcaatgagt	atatcttctg	tatatctctt	ccacttttga	ggattattctg
93241	aaataataca	acctttaata	taagttttctg	aaagtctgtt	agactttatat	atcggtcata
93301	ttactccaac	ttttgtacca	tttatttgat	tgatgggtaa	ttgtattctg	tttatattgtt
93361	tacaataata	ccaaatttat	aggaattgtgt	gagatgatgc	ctagcacatc	tgaaggtccc
93421	caaaaaatgt	agctttttat	tactactact	tcaaaagtga	caaagagagt	caactttata
93481	ctcttagagc	cccttaagaa	tatatgataa	tggtgcacaa	catgaacatt	gtggggacga
93541	cagtgtaacc	ctgtccacag	tcttaggaaa	tcttgttact	ccagtgtgct	agagtttctc
93601	cgacaactga	aaactgttgc	ctttattcaa	tttgtctcag	tagctctcca	cctcatgaag
93661	aaaaaagtca	tgttaggaga	gaagaagaag	atcaaaagaga	agcattttggg	cgagagattg
93721	cgggagacat	gtaactactg	caagatgtgt	agagaagggt	ttatatatttt	aaataatgtt
93781	caaatattga	tcatgaccoca	gaacttttcc	ttttgttctg	ggatacaggt	agatgaagga
93841	gagtcacatt	tgactttctca	tttgcccttc	ctttgtatatt	tcttttgttaa	tacttatcat
93901	taaatagaa	ttccaccggt	actcaagcca	ccccattgtg	gtatcagcta	tttaatttat
93961	ataacacatg	tgctaaaagt	tctgagcttt	tccgaaaact	aaaactctctc	ctgactccatg
94021	ttctgaaagg	acagtaagaa	aaggcacttt	acacatttta	ctgaacattct	aggtgtttag
94081	agagaccaga	caaatcacag	aagcactgtg	cacaccctca	gtaaagggaa	ttaggaaagaa
94141	cagtgtagctg	cagcttttcca	tatttaaagg	ttgtgtggaa	atcagaagaac	gactatttga
94201	ctactctagt	aaactccaaa	gtatatagtg	actcagacag	ggtaaggttta	tcttcatata
94261	tttaaatgtg	aaatagatag	ttttattctg	ccagtgtctc	tttcccaata	tgctattcag
94321	ggactttggc	ccgttgatct	atgtgtctct	gtgatctcca	agtccttatt	tcccaggtta
94381	ccaagtittgt	ctgtctcaat	cctgaagaag	aaagagcatg	gtgggtctag	accactgtta
94441	caatattgtta	cattcaagcc	tgagaatagc	acatctcact	tcaactcaca	cccgattggc
94501	aaatattctga	taactgcaaa	ggatactggc	aaatgcagtc	tagctgtgtg	cccagaaaga
94561	agagtaagta	gatttagtta	actgttagcc	acacacctgt	gggaaaaaat	ctaataaaaa
94621	atcatacttt	aaaatttttaa	ctgagaataa	cacacctcac	tttcaactcac	agttccattgg
94681	caagtatcca	gtaactgcga	aggatactgg	gaaatgtagt	ctggctgtgt	gccgaagacg
94741	aagagttaagt	agatttagtc	aactgttagc	cacacacctg	ctggaaaaat	gctaataaaa
94801	aatcatactt	taaaattttta	actgagaata	gcacacctca	ctttcaactca	cagctccattg
94861	gtagaatttc	agtaactgag	aaggatactg	ggaaatgtag	cttgctgtctg	gtgccagactg
94921	gaagagttaag	tagattttagt	caactgttag	ccacacatct	gttggaacaa	cgctaataaa
94981	aaaaatcata	tttaaaattt	taactgtctc	tttgagaagc	attttgcctt	tttaaaattat
95041	ataaaataata	actgtattcca	tattttatat	acagggtata	atttaataat	ctcacaataa
95101	gttcatattg	aaagagtcaa	tctaagaatt	catcactgac	ccacactggg	cttgaaatac
95161	aatctattac	tagctggaca	atcctatcag	atatcatctta	tttggaacag	ctctgtcaga
95221	taattctctca	gtattccagt	gtactgtttg	gtataacctta	tagataaagc	agagtggtct
95281	ttaacttgta	tcatcttgtt	catttcatata	gaacttaact	ataatagatga	ttgtgtctgt
95341	aatcccagca	ctctgggggg	ctgaggtggg	aggattactt	gagaccagga	gtttgagacc
95401	atctcgggca	acataggggag	accctgtctc	taaaaaataa	aaataaat ta	gccaggcatc
95461	gtctgcgaca	ctgtagtctc	cagctactct	gtaaagctga	atRcaggagg	ctacttgagc
95521	ctaggatgtc	gaggtctgag	tgaggtatga	tcataccatt	gcaactccagc	ctgagcaaga
95581	gagcaagaac	cccatctctg	aaaaataaaa	aaaaaagaac	ttaactaata	gacatttata
95641	agggcccccac	tacagactgt	tcacacctat	gtttataact	gacactaaag	cttgaacaga
95701	gtgtgcactgg	acagaattctc	aggtgttttag	aattgcgaag	tgaattgttg	tttaatttaa
95761	tgcttgaagg	taaatgtcag	ttttcatgca	tataatacat	tatttataac	tatagctacc
95821	atgctgtctct	ttaggtctcc	aaaaactatt	tatctataaa	tcaaaatttt	gtaccttttt
95881	actaatatct	ctccattttct	tccaccctga	ccctgggttaa	ccgaattttct	actttatttt
95941	tgtagctttt	ttttagattc	cacatataag	tgagataaag	aggtattttgt	tttccgtgtg
96001	ctggcttatt	tcaacttgaaa	tttgtaaga	gagaagattt	taagtattat	aagcacagca
96061	taaaaggggg	gtaattttct	gaagtgtatgc	ctatgtgtat	agtcatttca	gtctatatac
96121	atatatcaaa	acatcatgtt	gtacacctgt	aatatataca	atttttgttt	gtcaattata
96181	cctcaataaa	gttgataaaa	gatcaattgt	atttaacaga	catacatata	aaagacaaaa

```

96241 tcatccatca caataatgtt ttccatattt tcatagtatg cacagtctaa acaaggacct
96301 tcaacagatt ttaattttgt atttgttggg accattcaaa cagttttccat ctgggttatta
96361 taatccaagc tatgcattga gtgctcactg gaataaaaaa tgaattgttt ttgatccact
96421 gtcattcaaa cttttttcat gaagtaactg gaaaattgca tatattatcc tgtcaccccc
96481 aaatttatag ttcctgatgc ttttataaat ctatggctaa gggagacactg cagaattcaga
96541 aaattccctc tgggtgggtg tgggcttggg aacaactctaa gtacctgactg gaagctatta
96601 ctgaaagtgg ttctccccca gctgagaaat gtttcaaat taaagattta aaagtctccca
96661 accaattcac attgacatgtc agatgagtgt gtttaaagagc gtttaaaacta ttctttaaaa
96721 ttctaacttc atctctatat ggaagaagaa aagttgggag ttgttttctt aattgcccaca
96781 ctgtctaggt gctctccat gaaggccaat cagttagttt ctatgaagca atagaccacta
96841 agctctggcc tctgcattaa caagtgatca tttagtaaga gtccacaggaa agcagaatta
96901 aggataaaga caaaatctca ggtggaaatg tgggggaaag gtgtctctag taatttcttt
96961 ctcatgtYat agagaagaat gagcaagaat atagacaaag atcgaaatgg aatgtgtctg
97021 ctacagactc gttagactat aaaaattagc ttcatgtata cgtgcttaaa agcagagagta
97081 gggattcaaaa tattttacca ctggtatggt ttggggacca accctctagt acagctcctg
97141 agtctaagca accgatcggg caatgagtac cagcagtggg cgccctactg aagaggccaa
97201 ttgagcactc agtgtggcaa gagaagaattg gaggaaaaaa ggggctactg catggccaga
97261 ctattttctg agtaaaaaac actcaaaagt gatttctgtc ttgtctgtgc atggccaaac
97321 ctgattgatgc atcaggagat gcaggccagg tgaataaagg ccaactcagg ctatttgatg
97381 cctgtgtatc ttgtatttat cagagtcagc cagaggactc atgtcatctc ttgtgttata
97441 tccaaaaYgc tggcgtttgc tggactacta gactcttaga atgtcgagca aatgattgta
97501 atgtgctctc cctcaacttc ttccaactc ctaattgtct gtaatttttc caatagtaat
97561 atcatcaYct gaacttgcat tattttgtca cttttctccc tcaactaaaa tataagctta
97621 atgagaaagc agctttgttt tattattggt ttgtttcact ttgttctccc taattgcttc
97681 gaccatgtct ggtacatagt agactctcaa taaatctcgg ttgatttttg aataagtaat
97741 aacagtactc gtcacctagt aagggcttaa atctttcag agtaggcttc ttaattgca
97801 ttactgtctt ctttgtagcc ttaattctct taaaagtttt tgacacagtaa tggagtgaac
97861 ttaacaatag aactctaaat atccccctgc tccactcttg gatttccata ggggtccctc
97921 catctcagca caaactgacc gcaagctcaa atcaacctga ggcaaaactt tctaattatt
97981 ttgtttctta ttgtcttttt tcatatggtc agcccccggt gccctgtaag ccataggaaat
98041 gtgctctcca ctgccacagt ggaagtggtt atcaaaaaca aaattggcac aggcacaaagt
98101 tcaacaaaga aatttttatt caaagctatt gcaactaagg agagaggcca agatlaaagt
98161 tgaacttaac tgaactgagc tgaagggaag aactcaagct atctatgttt gctaattggc
98221 aatattcaaa gagggaacac tcaagtgctc cgtgtttgtt aattgctttt acccaaaaga
98281 aaagtaaatg tgcacttatc ttcatgtcaa tagacagttt tacaatgtgg agcaagatac
98341 ccacaaaagc tagcatctg tctcaccaca gagactgagc gagaggaggt ctatcttctc
98401 tgatgatagg atagccatca aagggaatgg tcccaggctc ttaagaaaga tactctgagg
98461 ttataaacgc tggccagaa cctttaaaat gatataatc tcaaaaggagc agagaaaggaa
98521 ttgtaaatca caggctttct aaagtaaatg ttctaagaaa agcagaggtca gggccctgga
98581 gtgtctaaag ttcatcatg ctgaaggctg ccttcgtcag ttttcaacct gggacacacc
98641 agtgacggcg tgtctcttgc tactctcaga gccaaaggctt attaagatgc cacagcagca
98701 ctgtgttcca tgaaaaagga gaaYgccctt ctctcccaaa gggtaacctg cagttacatg
98761 aacctttaag tagcactttc tgtgtaattc tagttcccat ctctctgtcc actggatcca
98821 atatcaaaat aaggactag aacatggcac agagtagtgc atctctccc tctccctcat
98881 ccaacttacc ttctcctcta actgcttaat tgtcaatcct gaaaactagt ttctcttact
98941 tttagcaaaa

```

SNW/ genomic sequence (SEQ ID NO: 3)

>14:76161051-76258200

```

1 aactcaagt atctctctgt gttggcctcc caaagcaatg ggattagagg tatatgccac
61 tgtgctctgg caaagcaagg agtattaaact ctgggggttaa tcttagctgt actgcttact
121 tgatctgggc atttttagat agtcattttaa tctcttagag cctcagtttc tctctatat
181 agtaacagta gactcaccca aaattaaagg aaaaatggWaa taagaaaaact taagcaggct
241 aaatttggct tcaaaaacta ctttttctta tgacatcata cttttccaac ttcttatggg
301 gctgaatcat gaaaggtgct tgggtgactg aaatatcttg agtcacaaag ctatttactt
361 cagttagggc taacagtaat gcttttgaga cacttgatga gattgtggtat ttaaacagag
421 ctataaggaa caccagctg gacagggtga gaaaaatgaa agagaaggcc ggtcgagatg
481 ctgcacgcct gtaatccagc cactttggga ggtcgaggca gcggtatcac gaagtcagga
541 gtccagagcc aacctggcca actatggtgaa accccctctc tactaaaaat acaaaataat
601 gttgggtgta gtggcgggcg cctgaaatcc cagctacgag aggctggggc accagaattg

```

661 cttcaacctg ggaggtgaaa gttgcagtga gctgagattg tgcactgta ctcacgcta  
721 ggaacagagc aagactccgt aaaaaaanaa aaaaaaanaa ataaaaagaa ataaaaagaa  
781 agtgcgtgag acagacagaa aaaaatgtgc agtgagacac tgggaatcct gtcgctaaa  
841 agtaaaagct aaaaatcct gttgcccct tctcagatga aatgtccaac tgcagctaca  
901 actctgacct ctgttaatta cttagaaatg aatttagtat ttcaggggct gaagtttctt  
961 ttgtatgtga gatatctaag gatggctcat tacattggtt aagagaccda cttcgaagtc  
1021 attgaagtcg aaacattaaa ggaaaaaaac ttattctctt cctgaatagt aaactcaggg  
1081 tctgcctgcg cttaaaatat atgaaaacagg ccgagcggtg gctcacacct gtaattccaa  
1141 caactttggg agccaaggca gtagatcac ctgaggtcag gaggttgggt ccaagactgc  
1201 caacatggcg aaaccccgct tctactaaaa atacaaaaa tgcacagca tgggtggcgc  
1261 tgcctgtgag cccagctact tgggagggct aagcagaaga atcatttgaa cctagagat  
1321 ggaggttgca gtgagccaag atcacgccac tgcacttcag cctgggtgac acagtgagac  
1381 gctgtctcag taaaaaaaag tatatatatt atatatcac acacatatat acacacacaY  
1441 aYatatatat atatgaaca ttctctttga aagagaatga tcttttatct tatcaactc  
1501 tccaagtgtg gagatttctg gacaaaaatc caaaagtgtt ctgatgtata cccagtctct  
1561 ataataaaaag gctatatatga aacctctgca cagtactatc agtgagttt tccagtacc  
1621 taggcactgt taaaacacct ttaagcattt tgcactgatt tcttttaata gccatcatt  
1681 atactgcact ataattcaac acactctcag aggtgatgtt agtgaactgt tctatcata  
1741 tgctaaaact aacgacacca ctgagctttt ataaaaggtt tctcttttcc tcccttgcca  
1801 gactgttaac caccactttc cccaccagtc tcatcagctg gcttgtaata tctaacttc  
1861 ttaattacta ttttaattctg ctttgagctg aggaactaaa caccocaaat tctatgact  
1921 caacacatac gtacttagt actgttttag gtccgtact agacacata tggctcattc  
1981 cctgatggct gacatttctaa aagcatatga catattatgg taaaattgat gtaaaaggag  
2041 ctgaaagagc gctggaagcg cctcgagag aaaaagtata ttaatttgg ggtgtgag  
2101 tctctgattt tgacaactct aagaccactt tttttcccta acaaatgaag gctgtgaaaa  
2161 gctgagctca atgatgcaat cattaacact taactatga gcatccaat tttatcagag  
2221 ctgttaggaa gcttggaagg aatgcctcat ctgacatctt gttggattga aagaaagaaa atataaagt  
2281 attaaaggag tgcatactct caacgttaaa gtggaattga aagaaagaaa atataaagt  
2341 gactacagtt ttggttcccc gttttccctc ttgctctgta tgttccctta aaaaaataat  
2401 tttttaaana tctatttact tgagaacagg ttatgagact ggctaatttt ttctattttt  
2461 ggccgaaacg ggggtcttgc catgccccag gctggtctcg aacccctggg ttcaagcgat  
2521 ccaccccgct cggctccgca agtgcgggat tacagcttgc agccacgctc cccgagctat  
2581 gttccccctt ttaattctca tggggnaaag Ytgagagaga gaacggtgct tgtattccat  
2641 atatacttgt ttgcatttta ttttcaaac ccgagggattc tcaagatgct tcaatgact  
2701 ccttctaaaa gagccagagg agaagacagg gactagtttt acgcacagtg gcttaacaaa  
2761 cagccgcgca tccctgtcag ttctgggatt cctcgggagt ggtttttttt aagctcgaga  
2821 agtctacctc accgtctaca aagactgtct atgggggtttt acaaaaaacc aaacagaaat  
2881 gtgagagac atttacagta agaataaaac acggattccc aggtctctttt gactgagcgc  
2941 ccccgccagg aatggccgcc taaagaacaa ggcattctcaa cgtcttatgt ctaagtgtga  
3001 cggttagtgg ggaattcgac tctgcagcca aaggaggtca ggaatcgttc aataacactt  
3061 gggtctggagg tgaagaagg cgaagaagac acgtaatgct cctcaagacg ccgggcaaa  
3121 cgatgtgagag acgcgcgcca ctctctctc ttgacttttt ggcacacagg cctttgcga  
3181 cgggtgctgct gggccgcgca gaagtgcagt acccttccca ggtctgctgg cccggccgg  
3241 ctctgactgct agaagcgga aagtttccga aaggcgctct cccgggctc agtgcgaga  
3301 ctgcccacag agccccagcg cgctgccact tccccactc ccgcccctat accctctgat  
3361 ccggaagcag attctctcgt gttccgggac cggaaatttt tccggggcc cgcacctgag  
3421 ctcgagaagg tgctttagtc tgaagatggc ggccctcaga cgaagaggtg ctgcggcgct  
3481 cgttgaatgt atcaatcagc cggttgcttt ttgagaaga attccttga ataggagac cctcgccgct  
3541 gagtgaagta tggagatgtg ggagttagtg aatttttagg tccaagtgat cctcaaaagt  
3601 tctgcttttt cgagggtact taagtcaagc ttggtgctga attagggag cctcccttga  
3661 catgcaagctg agcagaagaa aattttgatt gcagttaacc gtttagggat tgcctcttga  
3721 cttttggggt cagcaatgca cgggctcttg cctttaggag ttatttgggt taagatttag  
3781 ttaacttggtt accaagattt acaatttaga agggataatt ttgacttag tcttgaggat  
3841 ggaattcaagg tggcaagatt gtaggcgtgc agaggagcta aaaggcaatt agtgaactga  
3901 acttaaaacg tgaggacctg aattaggggc atggttagaag gatggtggtc atgtgactga  
3961 atgtttttaa agcaaatag atgcggtttc ttgatggctt agtggtgaga ttgttttaga  
4021 taaaaggaaa gggagaaatg atttctgggt ttttagcatt ggtgaaattg ttgttttaga  
4081 gttagagca gcccaagaga agatagtaga agtttgatg ttgtgaactt tttgtgactt  
4141 tgaggtttcc gtcaggcaat tggacaaaag tttgatatac ttgtggctga ttgtacact  
4201 ccgaagctac accgttgggg ttagaatttc aacttgcga tttaactt cttataact  
4261 gaataactga agttagtagt ggaaaaaggg ccacactgag gctcactgatt actcatMgat  
4321 aaaaaggag caataatacc caactcatag gatgtgtgag aggatttagt tatatgaga  
4381 actctctaga cttgtgttag atacagtga agtactcagc aaatgatat tattttccag  
4441 gtaactaaaa ttgaaactta cgggagaggc ctagtctaca gataaaggtt ggaataatta

4501 gttctgaaat ctttcaccac cttcaatttt tgtcagctgc agtttttgaga tggaggtttt  
4561 atattttgctt accaacatgt aaacatttcgt ggtcactctt ttgtatgata tttagtggtt  
4621 aggtcaacac ctggaacaaa aatggagcaa gatttagttt ctctcaagtt tttagtggtt  
4681 cagatgtgag tagtagaatg ggaataggtt tggttgtgct ccRgtttgtt aacaatgtgg  
4741 ttggtgaaga ccaagcccgag taaggtctcc aagtttccag gtgaacaktt atggaagggc  
4801 tggagagttt ccttagcatc ttgcggtgag aagtttccag acataaaggg agttttctct ttgtgttcca  
4861 aacgctttaa aagagtcacg attagtacaa gcttcgaaac agttttctct ttgtgttcca  
4921 ttgcttcagc aatatattat tttagtatct attctgtggt aggcatttgt ttctgtgctt  
4981 gcagctgttt ttcttccacc attctagacc tttagtgagc aggttttaact ttccagagctg  
5041 gttgtttttt aaatcagttg attctagacc tttaattact tttagtaaaa catatgttta  
5101 acattttttt gttctccagc atcatattct gtgagacott gatcttgcca gttggtttga  
5161 cttaacattt agagcccttc aaaaattagt ctctgtggat tkttggaagg attagtggaac  
5221 ttttaactat tcaattagcc aatattgagg cctctgctgt tatgtgtcac tgaatacata  
5281 tgggtttata tgcgtgagga gcttacaatt agtggtgggg atgaagagac ttagaacaaa  
5341 tatagccaat ttgtgaacca gtgaggtttt ggtgaactag gagaactaaa tagatatatt  
5401 aaaaagttaa tcagtgaaat aacagcactg tagttcagaa gagaacacaga tcgtgtagtt  
5461 tttagtcttt aagattatat actctaactt tttaggtctc cacagggctg gtacattgtc  
5521 ttgtactgtt ataggagccc aataaagtaa tactcagtg atctacttgg ttggtgctt  
5581 gaaggtagt tagtgttta caagttagcc agggggagat aggattacca acatattggag  
5641 gggagagaac aacatatgca aagaggttaa cccgagagc agataaagg attaacttaa  
5701 gttgtactag agagtaacta ggggttagca gtgggaaata agggcaggaag ggtaggtgag  
5761 ggcacactca ccatttttY taaaattgtga aagaccagta agttgggatt tatlgtggaa  
5821 attcttcagt taccgggaaa cacactggat tgtaatttgt cttttttctt ttgttgatct  
5881 gctcacaggt gctatagttc aactagcaca actagtcgga cttaaacact ogattattgc  
5941 tgaggtatga gactgcagcc aactacaaca ctYtctgtc catctcacca ccaagctctg  
6001 tacattatgt ctttatagtc agtaccatag ctgtctctca atctcttaagc tttaaggaa  
6061 atgcaaaagt gttcctgttc acagaanaat tggtagccct gttctttctac agtccaagat  
6121 agagaactca aatatagta actactttta atgtttttt ttgtccacaca agtcatcgta  
6181 aagaaacact tgcacagttc gggcatgtca gaaggtgcat tttaactttt taagattata  
6241 aggaanaagt gttgggaggt ggggtaggct aatgggtaca aaaaaaattg atgaataaag  
6301 acctatttga tagtacaaca gggtagactt ggtcaataat aattgtacat tttaaaaaaa  
6421 ctcaagatgt aattggatta ttgttaactc aaagataaaa tgcgtgaggg gtagtagacc  
6481 ccaataagta aaggataaat gcttgagggg atagataccc catgtctcat gatgtgcta  
6491 attttacattg catgctata tcaaaaacatc tcatgtactc cataataaat atacatatat  
6541 atatatataa atatatataa atatacacct aataatataa aaatatataa aaatatatac  
6601 ccaaaaataa atatacacct actatacacct cacaaaaatt gttSaaaaaa aagcttaagg  
6661 gtgttttaaaa aattcttatc ttgcaacagg gtctctgctc gtcccccctgc aggggtgcag  
6721 tggccttgatc tcaactgcagc ctctgcctcc caggcttaag tgactctctc ccttggcct  
6781 cctgagtatg tgggactata gacatgtgcc accatgccca gctaactttt ttgttatttt  
6841 tttagtggaga cagggtttgc ccatactgtc cacactggaa aattttttta atttcoattt  
6901 tctttaaagag acagggctct gtctgtcac cccaggtgga gtgcagtgat gcagctcagtg  
6961 cctcactgca cctctgcctc caggcttag gtgactctcc caccctagcc ccttagtagt  
7021 ctgggactca aggcagtcac caccacacct gggtaatttt tgtaattttt gatagagagg  
7081 ggtttctgca ttgtgcttat gctggctctg tcgaactcct tgcagccaca tatccgaca  
7141 cctcggcctc ccaaaattgt gagattacag gtgtgagcca ctgtctgtgc ttatttttgt  
7201 ggagatgggg gtcttgccat gttgccagg gtgacttga actctgtgct tcaagactat  
7261 ctctcgttat ggctcccaa agtgctggga ttacagcgt gagccactgt gcccgcccta  
7321 cagtcttat tgattacatt ttgagtata ctgtataaa aagttaaaaa gctttatcat  
7381 aggtatttgt gttagttttt ttattgtgtg ataacaaaat accacaagcR taagtgtcta  
7441 aaaaacacac atctctcaca gctgagcaca gtttagctgg gtctctctct taggactcta  
7501 catagctata tattccaggt gtacagctgg gctgtggtgt gtcctatcaa agggtaacat  
7561 ggggaanaag actgacttgc aagctccctc atgtgtgttg atgtgtgtct ttttttttt  
7621 ttttgagatg gagtctttct ccttgccta agctggagtt gcaagtgtgc agtctcagct  
7681 tactgcaccc tctgctcct gggttcaagc catctcctg ctactcctg cctagtcact  
7741 gggactacag gcatgcgcca ccacagccgg ctaatttttt tttttttttt ttttttgaga  
7801 cggagctcga cctctgtggc taggctggag taggtgtgag cgaactcggc taccaccaac  
7861 cctctgctcc tgggttcacg ccattctcct gctcagcct cctgagttag ttggaactca  
7921 ggcgcgcgac actacgcYcg gctaattttt tgattttta gtagagtag ttatYtgtat  
7981 gtttgccagg gctggtctcg aactcctgac tccaagtga tgcocgcgct gggcttccca  
8041 aaatgtcggg attacagagt tgagccttca tgcccgcca tgagctctca taatYtgat  
8101 gactgaggtc tgttttcttg ttgaccttg gctctgagtt gggagctcgt ctacagtagt  
8161 aactccctgc agttccttgt taMctagat atctgtaggg ttgtctacaa ttggcagct  
8221 tgcctcttta aagccagcaa gggctctcct agtctgcta agatggcata ttatcacaga  
8281 cataatcaca ggagtgacat cccattactt tYgctgtatt cttttgttta gaagcaagtc



8341	agggtttctcc	tMaccccaag	gaggggagga	gattatatag	gggttaacat	ggggaatgca
8401	tgagatcat	ggggcccac	tgaatctct	gctactaca	gtaattcta	aggagcatc
8461	gattctcaat	aattggcaga	ggatatttta	agttttatct	tcattgagta	ccattttctt
8521	ttctagattt	ataactaata	aaacaatat	tttatattaa	ctattttctc	ttgtgctttt
8581	tgtgaggtt	tagaatacag	cctataaatt	tatctctggc	tgtatttttc	tgtgggcttt
8641	tttaggtgt	gaagtggatc	ctagaagttc	acacacact	ctttgaaac	ctttttctct
8701	tgagcactg	taagggctgt	ggtaaaataa	aaacataaag	gaactctaat	tcttaattct
8761	tggtttaagt	ttacttgact	cccaagact	acagatttag	ttatataatg	tcgcggtgaa
8821	gtgttaggaa	atctggcact	caagttgttg	agaagttcaa	ctggtagact	caggtctgtt
8881	gtgtacttg	caaaatacca	aagttaagaa	tttatatttg	ctttgatcta	cggtgtactt
8941	ttctaagaat	ttatcctaca	gaaatgtgtg	gtcaagaata	ttccaaagca	gcattattta
9001	attctcaagt	agaagaagaa	acttaaatgt	tcatcactgg	gattgattag	attattaagt
9061	actgatgcaa	tggaatttta	tatagttgaa	tagaaaaagt	gggacactgt	attatggaca
9121	tgaaataatt	tctgtaagtt	tgaaaacttc	aattttttca	agagcattta	aataaaagt
9181	tccctgcmaa	aaaaatgaata	gatttttatt	ggtccataaa	taaaatagg	ctcgcgaaga
9241	atagactctt	gatttaaaaa	aaagggcgct	ttgtgtgttg	taaaaaaaa	taaaataga
9301	caaaaaaatt	ttttttttgt	tccattctatt	gtactctagt	tggaanaYct	gggttctctt
9361	ttgagacagg	tctctcNct	actcaaggct	ggagtgca	ggcacaata	agttcactgc
9421	aggcttgaac	tctctgggttc	aagggatctc	tctgctctgt	cctctcga	aggtgggtt
9481	agtcgtgtgc	caccacacct	gactaatttt	ttgtactctt	tgtagaattg	tgctctcacc
9541	attattgccca	ggctgggtctc	gagctcccg	actcaagcag	tactctcacc	tcagctccca
9601	aagtgtcgtt	attacaRga	tgagccaaca	caactggact	ggtattttt	ttttgagaca
9661	gccttgcctc	gtcccccagg	ctggagtgca	gagggctgat	ctcagctcac	tgcaatcgct
9721	gcctctcYgg	tccaagtgt	cccccttctc	cgctcaccga	agctctggg	actacagggc
9781	tgggccacaa	caccagacta	atttttgtat	tttttagtaga	gattggaatt	cacYatggt
9841	gccagctcgt	tccaaaactc	ctgacctcaa	gccactcacc	caactcagcc	ttccaaagKt
9901	ctgggagtga	gccactgcac	ctggcctatg	ttttaataa	aaacaaattt	tttaagctct
9961	tgactctgga	agtggtcaca	tgacaataaa	ggaatgcttt	agaaattgtt	tgatctacaa
10021	tgaatggatt	attgggacat	caccagaaac	tttccgtctt	ctaggttaat	ccctataaag
10081	ttatccttga	agtagaggag	caattgttaa	attttgtttt	ttttggcaca	gcacagactg
10141	acttgaattga	gactaattgag	gaccagtttg	tcatctacag	ggactagaca	ttctaatatc
10201	tgtttttctg	Ygaaggcgtg	gaggggcagg	gtacatccaa	gcaactcaga	attctgtttt
10261	cgtagtcgat	gtataggttg	aaacttttgc	acaagggaaa	atagcaaaat	tgttttaatt
10321	accagagaag	ctgtaactag	gttgagttaa	ggctaaatat	tattttttct	attttttNcat
10381	taggaacacg	tgatgatagt	cggtgtgtgt	gtgaggggtg	ctggattcag	aaatgagtg
10441	aaacaaggct	tgaggcttaa	ttttctagtt	tctagaacgc	aaagcttaga	caagagtact
10501	gatcaaaaat	tcctttagct	attactgtta	gtatggtttg	catctagaat	gacacagaaa
10561	gctcatgacc	cgatctctct	ctaattggcg	caaggaaaac	agcctcagac	acaacacagg
10621	gatttagctct	tggttaatca	ggaagggtgt	aaagcaaggt	tactgacact	tataaatgtc
10681	tgtagctcga	ttttYgcgtg	ggcatggttg	cttatgctctg	taatttcaSY	actcttgggg
10741	gccagatagc	gaagattgct	tgagcccagg	agtttgagac	cgccctgggt	aaagatggaa
10801	gaccccatct	gtcaaaaaaa	tataaaaaat	gagggcagaca	tggttggtatg	tacctgtgtg
10861	cccagctact	gaagaggctg	aggtggggag	attgcttgaa	cccagggaag	tgaggcttga
10921	gtgagttcgc	ttcataccac	tgacccccag	ctgggttgac	agagcaaac	ctgcctcaaa
10981	aataaaaaat	aattaaat	aaaagtcaga	tttttgagca	atgaaaaagt	tgttagtttg
11041	attttgaaac	catggaaact	tcttgaaagt	catgattaat	tttttctaat	tttttgaatt
11101	tttaggacaa	ggagactggc	tttcacagag	gtttgggttg	ggttcagttt	tcttcagaag
11161	aggagactcg	gaatgcacta	caacaggaaa	atcatattat	agatggagat	tgatgaattt
11221	tattttctat	ccagatatac	acatgatata	catgtaggtt	ctattttaatt	atgtatacat
11281	taaatagatc	ataaatgttg	aatgatgggg	acttgagag	atttgttaact	gaagctgcct
11341	taatgcttgt	taagaattgt	gggttgagg	ctgggcacgg	tggtctacgc	cttacttccc
11401	agcacttttg	gagggcgagg	cgggcgagtc	acRaggtcag	gagatcagga	ccatctctgc
11461	taacacggtg	aaaccccgct	ctactaaaa	atacaaaaaa	ttagcctggc	aaagttggcg
11521	gcgcctgttg	tccagctac	ctggggagct	gagggcaggag	aattgcctga	accocaggag
11581	gtggagcctg	cagtgagccg	agattgcgcc	actgcactcc	agcctggggc	atagcgagac
11641	tccatctcaa	acaaacaaac	aaacaaaaaa	acaaatggty	agaaatggty	cygttggtagg
11701	ccaggcgag	tggttcatgc	ctataatccc	agcacttttg	gagggcagg	cgggggggat
11761	caccagaggt	caggagtttg	agacacagct	gaccacacatg	gagaaacccc	attctctaat
11821	aaaaatacaa	aaaaattaga	tgggcggtgt	ggcacatgcc	tgtagtccca	gctcatcagg
11881	agactgagca	agagagaacc	agagacaga	ggttgcaagt	aactaaaaa	ggcccatgtc
11941	actccagcct	Sggcaacaag	agcaaaactc	catctcaaaa	aaaaaaaaaa	aaaaaaatg
12001	gggggttgta	gagtgccag	ttttaaaagg	tcagagaaag	tacaagatga	ccctttaagt
12061	ggtagcaga	tatgtaaac	agtgacata	acagcttact	ttctttttt	ctttttgaga
12121	gtcttgcctc	gtgtgccag	gctagagtc	aatggtgcaa	tcttggtaca	ctgcacccct

12181 gcectcctgg gttcaagcga ttctcctggc tcaacctccc gagtagctgg gattacaggc  
12241 atgcgcaccc agcgaactgc aattttttttg gatttttttag tagtagctgg gtttctcatc  
12301 gtttgctcagg ctggctctcaa actcctggacc tcagggaactc cgccctccact ggcctcccaa  
12361 agtgcgggga ttacaggcgt gagccacagc gctctggccag cagctctgct ttcaagtgtag  
12421 cgaagctggg gctgaactct tggcttaacc accatgtgat cctggggccaa actgatttac  
12481 tttttttttt taaatttttag tttaatttag acagggctct gctatgtgtg actggtgtgt  
12541 tgaactccct gctcgaagca atcctctgttc ctgagcttgc caaaatgtgt gattacaggc  
12601 cctgagccac tgccctctgtt ctgattttaca atttttaatt gagggtagg agcataaat  
12661 accaacaaga atgttaaaaa gaatttttagg agaaaaYgaa agaaccaagg accaaaaaaa  
12721 actctcggg ttattcattat tcaactagca ctccctaata aagtgttaat acttttgtaa  
12781 gtttttggag acattgacct tctctccctta cagtgcttat tttttttaagg tccaggttca  
12841 cactagaagg ccaaaacttc cgcaaacatc tgatgatgaa aagaagaatt tttagactg  
12901 cagcctatta ataaagttaa cataactgag aattttgtct aaatgttttt atttgaaaaa  
12961 aatagtgtga ccaagcaaga gKttactttg cccactccaa attaaaaacag agcacatag  
13021 gggcaaaaatt tatttggcag gacagttcca gtatgtgaac atctctctcc tcaactgtgt  
13081 tgggtgaact ttactcatat gcagctgttc ttacacaaa cattaacccc aaactctag  
13141 tgtcacataa aagttaagtgg tcttgacttt gtatgtgggg cagcatgttc tataaattgt  
13201 gaaggtgggg agaaagcaca acacaaccca ctcttataaa aaaaactaaat aattcaaat  
13261 agaattttct atccccccca ttctccaggt aataaaaaagt agtctgggga gctggcacc  
13321 aatagtgtgt ttatctctga ccattttaca agtgttcccc catatgactt cctgtcatat  
13381 ggttatgggt aagagttcat tcaatttggga gagacgtgtg cctattctct cctctctct  
13441 ttgctctcat gctcgtgttc ctggggcggg ctgtatctg agggcttttt agagccacca  
13501 tgccttctgg ctcttcccaa aaactgttcc aaacccaaag gatctctctc aaactgcact  
13561 ggtcctcttc ggccctctctg tctactggtc gaaccagaaa actcctgttc gggaacaaat  
13621 ctaaggtaaa ggagaaaaaa ctatgaacta ctttggcttc accaggttaa acactgaat  
13681 tgaagtgtga ttcttggcac ctgttggctc ttattctggc ttctagtgta caccataca  
13741 tgcctctgtc cagattttta ctgggctat aaatactctt ggccatattct ttaccactct  
13801 tccaggcggt atcataaaca ttataaattt catcttctcc accctgaatt cagctgtcca  
13861 taccctgttg aaaaaaggat gacattaaat aaaagtgtaa acattgttta caaaaagctg  
13921 aatcataaga tgtttacagt aaacctctgc taagtacag ctacactttt caaactctg  
13981 aactgcaatt taaaaagctt tcatataaaa actgttagca catgaacttc ataaacttt  
14041 tggattcttt tttttttttt ttcaagacag agtctcaact tgcacacag ctgtggagtgc  
14101 agtgggtgga tcttggctca ctgcagccac cgctcccaag tccagtgat cctctgtct  
14161 cgcctcccca ggtagctggg attacagcca tgtgccacca caccacagta atttttgtat  
14221 ttttagtgga gacagagttt caccatgttg gccaggtctg tctgaaccc ctgacctcag  
14281 gtgactgcct gccttggctt cacaagtgct tacatagctc actggaagct tgacctctct  
14341 gtgctcaggt gtctctccca cctcagcttc ccaagtagct gggagtacag gtgtgtgcca  
14401 ccatgtccag ctaattttYt gttagacagg ttttgccata ttgccaggc cggtctttaa  
14461 ctgggttcaa gcaatctgcc tgcttaggtc tcacaaagt Ytaggattac aggcgtgagc  
14521 cacagtgccc agccagaatt attttcatca caatttttaa atgttttaag cttaagcttc  
14581 taacaggtga ttactttgaa aatatttact tacgtgaaat ttcaacccca gatccctttg  
14641 cctagctgtg aagttggcat tgtatagtgg ttaagaaagc ttgaattatg gctcgccaga  
14701 tgggtccagc cctgtaattc cagcactttg ggaggccaag gcagagtagt cagaggtca  
14761 ggagtgcaag acccgcttga cccaacgtga tgaaacctg tctctactaa ggatacaaaa  
14821 attagctggc catgttgcca catgcccgtt gataactcca gctactcagg aggcgtgagc  
14881 agaggaatcc ctgaaacctg ggaggtggag gtgtcagtg gccagagtgg caccactga  
14941 ctccagcctg ggcagacag tgaggactcca tatcagaaaa caMaacaaaa caaaaaaaac  
15001 ttgaaatca agacacccag ttgaaactt aattccatca gtcacagct aaactgtgt  
15061 tcaatttcc tcatctataa cataagagat aatagtata ttctatagga tttttgttag  
15121 aatttagata ataaacagac ttagaacaaat gccaaagata taaatgcat ctaataaatt  
15181 atcaactaga agagcttact taaaaacctc aagtcataat taagtctccc actctactt  
15241 cagtagaagc ttacacttct gtattcttgt agtcttctgt ggtgctggcaY ggcgccggt  
15301 aggcctaaag tgcctagtt agctaatggc aattgtactt agaaatcaga ttctctgaca  
15361 ccagcctag tactttaac cgttttatac gcttacttt taaaagata ccaaatgtat  
15421 attttctat agcattaaaa gaaaaatcca aaatcaaca acaaaaaaac ccaRaggctg  
15481 cttttttagg acactgact ttttttttt ttgagtaga gctcagta gccagctag  
15541 agtgcaagtg cacaactctg gctcaccagg gctagcttct caattctctg gctcagctg  
15601 gggtatttaca ggcgtgcacc accacacctg gctagtttct gatttttag tagagtagt  
15661 gttccacatt gttggccagg ctggttgtga actctcgacc tcaagtgatc tgcccgtctc  
15721 ggccctccaaa tgtgttggga ttacaggcaa aagccacca gctgggctc cctggttcta  
15781 ttgtgaaac aaaaagttga cataccacaa caccagttt ttaactcact attttcttc  
15841 ttctcctagt gctcacaatc taagcccaat ataagaagt taaactcaaa cttttctcat  
15901 tgtctcaag ttccaacata caacttccaa tttttttgtc atcgtctgg caactacatc  
15961 caaatgtaaa gtatgWtaat ttctgttttc ctgagagtag gttatccatc caaacacac

16021 aatatgctgt tcttaaaactt gtaccttggga ttgggtgaag agcctttgggt catactgaac  
16081 ttcattgtgga gtccgaggat taggaacacc gagagcaata acttcaactga tatcccgatt  
16141 ttcattctctt tgaagtttcg acctattttg aaatacagaca tcaactaactg aaactctttt  
16201 atagacaagc tgaatatgtt caatagaaaat acacaacatc tcagaaaaacg aggagtgtga  
16261 tatgttcacat tcaatctctg tgaacatata ttcccttgatg atacatMatc aaagtgtggag  
16321 cccctataaaa gcatttcaga tttaggggag agttttacaa tttagagctgc ctgttagctc  
16381 aaaaagaacat ctaacagatt cgttgacatg gaattcttta atataagcaa tcaacttaca  
16441 aMcatgctga tataaaaaa tgcattggct tatatttact atcatatat gatgggttgg  
16501 gggttgggga agtggaataca atttagatgt agcacacgaa acatttttaa taagtctaga  
16561 acgactctctt tatctaacca catttagtaa ttttatccca ctgaattttt gaggaaaaca  
16621 tgtcaatgac tgaatttaga tgatttttga cattcagatt aatttggttt cctggtctgt  
16681 gcattatata tgcaacctac atcccaatga aacctggta actgctatta ctgttttac  
16741 taaaaactca ggtcacaaaa tgattttcta aggctgtttg attttcatcc agtttacaatt  
16801 tcacagctctt tctatgttgt caattctgat agccacttta ctgcttttga gggcccaatt  
16861 tatccacaga gccattgttaa cttttacaa tttaggaata ctaaccaaga ctgtctataa  
16921 ttataaactt tcaacttttg ctccctcttc tacatgtact ttaccctcaa ctctcaagag  
16981 taagagcttt tgccttccca tgacatcatt tatccaggga gatgaataat gccaatataat  
17041 accagacgga tcttttttca tgettattta tttttatttt attttttatt tatctttatg  
17101 accagggaat tcaactctgc gccaggctgt gaattcgagt gcacagacttt ggcctccctga  
17161 aactctccgc tccagggttc aagtgtattc cccacttcgg cctctcgagt cctgttcgact  
17221 accggcatgt gccaccacac ccagctaatt ttgtYatttt tagtagagat ggggtttcac  
17381 catgtttggc aggtctgtct tgaactcctg acctcaagt attcgcgtgc ttgtgctctg  
17341 caaagtgtct ggattacagg tgtcagccac tgtgcccgcc ctcttttttg tgtttaaatt  
17401 atgactttgaa aaatgtatta ggctcatggg atgacagag cacagaaaaa tgttcttatg  
17461 ggaagaaaaat attctataat tgcatttaac cctcaactca aaattccact taagtattga  
17521 gcgttccctc tacttttcag aacttttagt cctcagcctt tgttgagcca tagactccc  
17581 caattctcat ttactctat gccctcatat atataattag ttgttcttat acgtttatca  
17641 ttactaaag tgctttgcaa atttttccaa attattttat ttgttagctat cttgtttctg  
17701 actgtaagtt aaactactcc taaaaatgag ctgattttct acttttttca aaactgcctc  
17761 tctctccaat tgcactagat gcagtactat gcatataaaa tgaattccaat aaagcatga  
17821 aagcttatttt cctgtctaaa gaggcacata atccatttat ctcaactatt caggtagaaa  
17881 aagaacggac tataaacagtc caaatttact actacactct ttccaaaaat ggactctga  
17941 ttccgttttc tatattccca tcatttccaa ttctgactta agtgatggaa gattaattcc  
18001 aaagacagtc aactagagct aagacaaaat agttaagtac aagaaaattt agtctaatt  
18061 tattaagtta tcaactaaaa gaagaaaggg agagaagcaa aactgatctt agtggactaa  
18121 aactctgctc taaaaggtag aaaaaggcca ctcttcata ttgtgtacata tctttttttt  
18181 tttttttttt ttttgtgaga tggagtctcg ctctgtcgcc caggctggaa tgcagtgcg  
18241 tgatctcgcc tcaactgcaa ctccactctc cgggttcacg ccatctcctc gccctagcct  
18301 gccaaagtgc tgggactaca ggcgcctgcc acgacatgc ccggtctaatt ttttgtact  
18361 tttagtagaca cagggtttca tegtgttagc caggatggtc tgactctctc gacctctgg  
18421 tccgcccgca ttggcctccc aaagtgtctg gattacaggg gtgaaccact ccgcccgccc  
18481 accatgatgc gtacacatct taaataaacac ctactctta tcaaggagctg cgcctgaag  
18541 attccggtga tgcgtctctc ctttgcctc gtcatcgccg atttcatctc tctcaactg  
18601 ctcccactcc tctgtgtgaa cagttgaaaa gtaactcact ttaatatgtc ttatagtgt  
18661 gcatcagctg acagctgtgt atatatataa tttttaaaaa agagacagca tctcaactgc  
18721 tcaaccagtc tggagtgtag tagcatgaac atgggtcact cgagccacta cctctggat  
18781 tcaagtgtac ctctgcctc atcctcccca gtacctggga ccaagatgt gtgccacat  
18841 gcctgtgata ttttttaatt ttttgtaaag atggggctcc acttatgttg tcaaggctgg  
18901 tctcaaaact caggctctag tgatcttccc acctaggcct cccaaaagKgc tggagtca  
18961 ggtgtgagcc actgcactca gccccattag tattttgttt gttttgagac ttatgtctgc  
19021 tctgtctgcc caggctggaa tgcagtggcg cgactcagc tcaaccgcaac cctctgctcc  
19081 cgggttcagg cgattctctc gcctcagcct cccaagtgc tgggatata ctagctgacc  
19141 accatggcgg gtcaattttt tattttttagt agagaatgag ttcaaccaag ctgattctaa  
19201 actcctgacc tcaagtgtac tgcgcacccc agccaattat tgcgaaccca ctgaattttt  
19261 taaccttaga aaattccaaa ctcttttga tgcacactca gcccttttaa ccttacttgc  
19321 aattaaattta agtcaataaa atgatatagg tttaattttt gaatgtgaat tctagaata  
19381 agcatggcgg taaaataatt aagggtaaaa tgtattgctg tcttcaactt acttWgaaat  
19441 gcaatgcagc taactcgaca gacagatgga cagaaaatg ggcaaaatg tgaataagca  
19501 aatttatcaa tgttaattgc ttttaaaatt taattttgat atataatcat tatacattat  
19561 tatgggggat atatatgtga tacaatgaaca aaatgtgtaa taatcaaat agggtaatt  
19621 ggtatcccat caccctaaac atttgtttgt gttgggaaca ttctaaatct tgcactatt  
19681 ttgcaatata caataaatta tcaataacta ttctctctc ctactagaa cattagaact  
19741 aataaattat caYtaactat tcaaccactc gtactaggaa ctttagaact cactctccc  
19801 aactgtattt ctgtaccatc aaccaactct cttcatctcc cttcaactcc

19861 aaccactggt aaccaccatt ctactctctg ctctcttgag accaactttt ttagctccca  
19921 cacaacagtg ggaagattca gttactgttt ttctgtgcgc ggcctattaca gcttaataacc  
19981 ttacagttcca tccatattgt tgcataatcc atgattttcat YMTttttttat ggcctagatga  
20041 tatccattctg ttatataaga ccaacttttc ttlatccatt caccacttgg accactcttg  
20101 attccgtatc ttggttaactg taaagagcac tgcataaacc atggaaagtgc agatatccct  
20161 ttctgttttc atatactgat ttctttttgg atataacc ccctctctga ttctgggaatc  
20221 cgtatgtggt ctattttccag tttttttgaa aactgccata ctgtttccca ttctggcctg  
20281 actaattttt attcttacta aaaaatgact agtgttcccc cccctctctga tccctgcacg  
20341 catctattat tgccttttta atgataacca ttttaactgg agtgacataa tatctactg  
20401 tggttttgat ttgaatttcc ctgatgatta gtgatgctga acatttttga tatcacctgt  
20461 tggccataag tatgtctctt tttagaacgc tctattcaga tcRtttgccc atttttaaat  
20521 cagattatct cgtttctttg cgattgagtt ctgtacatt ctggttatta attctttctt  
20581 aaatggataK tttagcaacca tttcttccca ttctgtaggt ggcccttcca gtctgttgac  
20641 tgctgtgcag aagctttttca gcttgacgca atccatttg tctatattg cttttgcctg  
20701 ctgtgctttt gagattttac ccaaaaaatc ttgtccaga ccaatgtctt gaagcatgtc  
20761 cccagttttt tcatctagca gtttcaatgt tcaggcctta cattttaagtc ttaatcatc  
20821 ttgtatttga ttctttgatg tgataagaaa taagaggtct attttcattc ttttgcattg  
20891 gactattcgg ttctcacgc attatgtatt gaagagacta ttctttcccc aaatttttct  
20941 ttgggcatca cggttaaaaa tccacgcat ttgctgggcat ggtgtgtggt catgctcttc  
21001 acgcatcttg ggagctgag cggggtggat cacttgaggc caggaggttg agaccacct  
21061 ggggtctaat gcaaaacccc gtctctaaaa taacacaaa attagccagg cgtgtgtgtg  
21121 cgtgcctaat tccagttac ttggggagcc aaggtgggag gatagctga tcttggagag  
21181 cggaggttgc agtgagctaa gattacacca ctgcactcca ggctggatga cagatggaga  
21241 cgttgcctcc aaacaatcaa aaattccatg catttatagc cgactccatt attcttggat  
21301 ttctctattct gttccactgg tctatgtgtc tgtttttatg gcagtaacct gctcgttttg  
21361 ttactatagc cttgtagtat attttgaagc caggcagtg gtgcctccca gtttctttg  
21421 tttttattca ggaattgctt ggttatttgg ggtcttttgg ggttccattc aaattttatg  
21481 attttttttt ctatttcttt gaagaatgtc attggtatt ttgtagggat tgcattgtgt  
21541 cttagagctag ctttgggtag tatgaacatt ttaacataa taattttttc aattcatgag  
21601 catgctcatc ctttccattt tttagtgttc tctacaatt ttctttttcc aaattttaat  
21661 tttttttttt taaaaagaga tagggtctta ctgtgttggc cagggtgtgc ttgaacacct  
21721 gacctcaagt aatccttctg ccttgccctc ccaaaagtgt gggatcacag gttgagcca  
21781 ccatgcccag cctcttcaat tacKttcatc agtcttttat cattttcttt gtggcgactc  
21841 ttactctttt tggttaaagt tattctaggg aggtctatgt cagtgactca cgtctgtaat  
21901 accagcactt tggggagcYg aggcaggcag atccactgag gctcgagacc agcctggcca  
21961 acgtggtgaa acccatctc tcttaaaaaW aaaaaaaatt agccacggct ggtgtgtgct  
22021 cctgtttgtc cttagctact agggaggcga gacacaagga ttgcttgagc ctgggaaagc  
22081 gaggttgacg tgtgctgaga tcatgccact gcactccagc ctggttgaca agagtgaac  
22141 tccgtctcaa aaaaaaaaaa cctttattcc taggtagttt ttggtagcta ttataaagg  
22201 gactgccttc ttgtttttca gcttgggtat cctatttgta tatagaaagt ctacagaatt  
22261 ttgtatgttg gttttgtatc ttgtaagtgt actgaattcg ttatcagtt tttaagtttt  
22321 tggctgcctc attaggtttt tctaaataa ggaatcact gctctgaca aaatttaatt  
22381 tgatgtcttg tttttttttg tttttgagat gRagtcttg ttgtttggcc aggtcggagt  
22441 gcatggcga gatctcagct cgctgtaacc ctgcctctc caggctcaagc aattcttggg  
22501 actcagctc ctgagtagct gggattacag gtgaaYgcca tcataccagc ctaattctg  
22561 taatttttat acagacgggg ttctcatgct ttggccaggc ttgtcttgaa cctctgacct  
22621 caagtgtatg cccattctcag cctcccaag ttgctgggtt acaggtgtga gccactgtg  
22681 ctggccttcc ttccaatct ggaatgcctt tattttctc ttttgccraa atttctctg  
22741 caaggacttc agtcccatg ttgagtaca gtggtgaag ttggcatcct tgccctctc  
22801 cagatctctg agtaaaagact tccaatgct ctccattcag tatctaggt tggggtttg  
22861 catatagtc cttaagtatt tttagttatg ttcttctaa gcccaacttt ttgagagt  
22921 ttatcatgaa gggatgttga attttttcaa atgcttttct attgtctact gaaatgagc  
22981 tatggttctt attcttggtt ctattaatgt gatgtatcac atttactgat ttgcataat  
23041 tgaacctatt ttgcactct gggatgaagg atagtccagt tgattatgat tgaatgata  
23101 ttttaagtgt ttgaaactg ttgttgtaga tttgttgaag gatttttgca tctaagatta  
23161 tcaggagatg tagcctgtgg tttctctttt ttgttataat ttgtctagat ttgtctcag  
23221 tataactagt accctgtaga atgaatttgg aagcatccc ttctcttcaa ttttttggaa  
23281 tagtctgagt agaactggta ttttagtctt taaatgtttg gaaaaattgc cagatgaagc  
23341 ttgcaaatct tgaactttc ttctatggga gcaatgttaa ttgttaaaaa aaaaaaaaaa  
23401 attagttag aaacttagtg gttagagata cRggtgttca tcatataatt cttcagact  
23461 ttctgtaKgt ttgaaaaatg tcaataataa atgtcaggaa aggctgggtg ttgtgctca  
23521 tgctctgata cccagcact ttgagaggct aggcagtgag atccactgct tgcaggagat  
23581 tgagaccagc ctggccaata ttgtaaaacc ccatctctac taataataca aaaaattagc  
23641 tgggcaggtg gatgggcacc tttaatccca gctacttga agccttaggc aggagaattg

23701 cttgaacccca ggaggcggag gttgcagtc gctgagattg catcattaca cttcagcctg  
23761 ggcaacaaga gYgaactcc gtttcaaaaa aaccaataaa taataataatg aaaaaagtca  
23821 gaaaaaagaa attatgcata atgtttccct tatattttca aagaattctat tgaataaaga  
23881 atctctcagct tatttccctt aggggctggc tggatcttga tttccactgc atttataatg  
23941 atcttctatt ttttctaccc tcgccagctc gagggaaatgg agtccagttat agagcgtaaag  
24001 atttggagca agcctgtctg gattcaaatc ctgtttctat agacataatg ccatgtgacc  
24061 tggcgagatt acttaatact atagccctta gtttctctct ctgtaaaaaa gtcacaaata  
24121 tagtatcaact acatagggtt atccagggga ttatgagata agtcaataaa gttactcaga  
24181 gtgtgtgccca ccaataagaag caccgtataa atgttagttt ttactactat ctggcaagag  
24241 atccttaata atttttaatt tcttgagaac gattctgttg tctttttctt acYactctag  
24301 agaactcaaa tgcagtaagt tttatYttcta attatcttct atggttaggt atacaaccag  
24361 gcaggtagtt atatctcacc ttattaaagtc attcataat aactgagggg agtggatcac  
24421 tacttaaaac cctgtgtatt tccctatggt tgaactctag actcttgttaa ccaattacct  
24481 ggcacctcca tcttaaacct aacatatcct taactataat aaattcttca tctttccctc  
24541 ccaaacctgt tccctctgca gtcttcccaa tattagtaa taaaagtaat attttctttt  
24601 atcctgagca attggcaatt taactcttcc caactgcta ggataaaaagt tctggagtcg  
24661 tccctaaact ctactttttt tctgtatctcc acctaatcca caaacaattt ttgtcaactc  
24721 tactcaaaa atacatccaa tatgggacca tgtttttgtt acctcttgtt acYactctac  
24781 ttacagtttta tttggaaggc attagcccat gccactcca cagttatctta ttttacaact  
24841 agcagccagg gaggctcttt aaaaatatat gttaggggaa aggtcatatc ggtcaaaagt  
24901 ctcaatttagt caggatgagt aagtctggga gatctaatat actgtatgtt aactataagt  
24961 aacagtagct catgtatatt ttgaaatttg ccaagagagt agatcttaag tgtctcacc  
25021 aaaaaatttta aaaagtaact ataaagagat actggttaag taattagctt ggcagtaggt  
25081 aatatcttcc aggttatatt aaaaatcat gttgtatgtt ttatacaatt tttattgttc  
25141 aattgtacct caataaagct ggaaaacagt atgtcaggtc attactctgc tcaaaaacct  
25201 gcaattgctt ccttttccac ttgtagtaaa agctttatct ttaggttaaa gggccatcca  
25261 tgaattgtcc actactcttc ccacctgccc taactattct gacacctcat ctctcatcta  
25321 tccctctctt tctcatcaaa atccagtcac attggcctgt cttctctccc ttgtggccca  
25381 ctaggcaca tcccatttca gagccatgta ctgtctggcc ctctctgtgt ggtgtctctt  
25441 ccccacgaaa tccacatggt ttgcttctct acttttctca agctattctg ctcaacagct  
25501 catctcagaa gtgagggcta tttgaccact ctccaccttt ggcaatttta tctctctt  
25561 cctgtatata ttttctctcc acagcaatta ttacctttta ctatcacaa tgatttaatt  
25621 ctgtctgtct acactcacca gaattgttag tgttataccc catcccccta cccactgtct  
25681 agcaaaaaaa ctttgtgtct aaaaagaaac ctggcacaca cagatgtctc tagtatgtc  
25741 aataaaatga aagaatgaat aaatgtacca ggtaagaatt tagtaatgaa aaaaaataa  
25801 caccctcac aaaaattctt gaattatact taggactcga cttcccactc taattgaatt  
25861 aatacccaa tgtagaccac taagttaagt cagagagaaa atactggagc atttaatgat  
25921 gagaatatga ttggagccta caagatagaa aaactctgga tttactcacc atatgagggg  
25981 gctgaattat tctgttagat acctaatctc tcttaagttg ttgagttaag ataaagaaa  
26041 atcaagccta tcatttgagt aataatgcaa ttttgaaaag tcaactgtac agaatgttaa  
26101 aaatttctgg tacagtccca aactgaagac ccaacttttt ctatctctct tttgtcttgg  
26161 ctactaagc cagagattac tctccatttc cactatata ctgatgtac gattttatag  
26221 ttatattgat ctactttata tgcagtagga aaaagatttt ctaaaataag ataYattttt  
26281 tccaccagc aaaaattcaat tcagtaatga gaatgtcata cctttttcca catagttgtt  
26341 gatccagctt cttctctccc tggcttctgt ggccattctc ctaagttttc ctctatgtt  
26401 tctctttctt tcttgagcca tttttctctc cgaatttcca cgaatttcca cagctcacg  
26461 agcctgttgg taaagtcaaa tgttaaacag gacactatca tgggataaat atttatgtct  
26521 atcatccacc tccctaactc ctggctggta agagcaatta tcagaaattc agattttgta  
26581 tgtctaatta aaacaaagag cttgtctcac taaaagataa cctgctctga agcacactga  
26641 aggtcagctt ttttctctgg gaattttta gtgattgaaa ctactgtgac ccaatttaat  
26701 ctgttaagct aagttatttc ctaatgcttt ctgcaaaaa atctaaagaa caaaatttaa  
26761 gccagccata gtgctctctt cctgcaactc tagcactttg ggagccagg ggtgaagag  
26821 tgtctgagct caggagcttg ggaccagctt aggcacata gtgagatcca tctcaaaaa  
26881 ataacatttt caaaattaat aacactttca aagttaatca agtcaactca agtcaaaatc  
26941 ttaacttaca tctgctagat aatttttata agggtactca ctgtgataca aaagcaatt  
27001 tttttatgtt ccaaacacat gctatgatgc atgtgaaact caatatata tatatttcc  
27061 ataattttta tttttacata ttttttttga gatgaagtgt ctactctctg tggccaggtc  
27121 gaagtgcagt ggtgcgttct tggctcactg caacctccac ctctgtctcg cgagagatc  
27181 tctgctctca gctcccaa gtagagatgg agtttcacca caatgtctga tgcacactcc  
27241 tgtattttta gtagagatgg agtttcacca caatgtctgg cttgtctctg actcttcaat  
27301 ctcaaggaat ccaccagctt tggcctccca aaatgtctgg attatcaca tgggccatgt  
27361 tggccagcct aattccataa tatttaaatc attatagatg ggcacagctt  
27421 atgaacgcca ttttatgtaS tattcaattc attagaatgt ttaactgctc tttctgcca  
27481 ctgccccttt gggggacatc tacaatcaaa atagagaaa atttgaatc aaaaattatg

27541 aaaccttagca atcttcaaaa aatgggatca tcttaggaaa ccattattctt gacatgttgt  
27601 gaggggcgcc ggtaccact gtgggggtga tgtcacacatc tacaaaaagct gggagctatgc  
27661 agagagctgc caactagcag ctgccatttc tgacacacgt accatgggtg agacttctct  
27721 ttgggtctgc gttctatagt taattatgca gtaccttagg taacacagaa tggattttaa  
27781 aagtaataagc agcatttaac aaattataac aaaccacact tccgatcagc aatgtatagg  
27841 gctcttcacca atttggcgaa attttcattt atgtgtactg tctgtatcgc tcttccatca  
27901 cgacgcagag gtttgtctaa tggaaattga taacctaga gtgaaagcagc acagaaaaacc  
27961 cagtcacaga agtgcttcaa ttaactgcct gaagttaaat taattttatt ttcttcgcgc  
28021 ttgagattctg ctctgtcacc ctggctggag ggcaggggtg ccacttcggc tcaactgcac  
28081 ccttgcctcc ttgggttcaa cgagtctcct acctcagcct ccacagtagc tgggaacctga  
28141 ggcgtgcgac acacactcgg ctagtttttc tatttttagt agacacggag ttccaccatg  
28201 ttggccaaag ttgttttgaa ttctgacct tgagtgtacc gcccctggcc tcccaagtgt  
28261 ctgggattac aggcattgagc catcgagcct ggctaaaaat aaattttaat attgccacaa  
28321 tttagaaccc acttaactgg aggtttcttt tctctgtcat ttcttccaca atgtgcata  
28381 acaacagcag ctatcaatga gtgctttatg tagaagatca agcaactcaac cctatggtct  
28441 catttaacct tcacgataac tctataaagc ttctaactat attgtgtcca ttctactgat  
28501 gctttgtgtaa catccccaag gtcatacagt tagtaagtgg gcgatttaga tttaaccact  
28561 Ncaggattgc agaacccaa ctctggaaaa atagatattt cgtttaataa ttctcgattt  
28621 taataaaaag aaactcactgt caaacagctc aaacagttgt cctcaagtaa atttaccagt  
28681 caaacaagaag tagaatcact ataactttag catctatagg aagtttaact tctgtactgt  
28741 tcttaagtat ctgttatatt cttagaatcc tgaattttta gtaaatagaa aaaaaagcct  
28801 ggggtcagta ctgcaagcct gtaactccca cactttggga gccaagcgca ggtggattgc  
28861 ttgagccagc gagttcaaga ccagcctggc aaacatggta aaccocgctc tcccaaaaa  
28921 tacaaaaagc tagccaggca ttgtggggca cactttgggt cccaactact cgggagctgt  
28981 aggtggggag atcgctgtgag cctggggagt tgaggctgca gtgagccatg atcataccac  
29041 tgtactccag cctgggcaac agagcgcagac cgtctcaaaa agaaaaaaaa aaaaaaaatg  
29101 aaccaaaaga aaaaaaaaaa aaatgttttg ctataacact gaattttctac attctaaaaa  
29161 ttgtactccc atttcaataa ctcaagaact gtacttgaat gctatgttaa ttctcaggct  
29221 tagctttatt ttccagatag ctgactagtc tctcaaatgt tgcattttag tcaatggctt  
29281 cataataaaa tgaagcatta tagaaccagc tcaagatacc acctagttac taaaaccaca  
29341 tcaactcaaa tcagaccagt cggcaccctc ttgttttagg gtgagtatat aaatttctct  
29401 aagaaacctg aaatggtctt ctttatccat caaaaaataa atcatgaact attgattcag  
29461 gaagatgaaa cagttaactag caaatctatc ccacatacta cagatttcta accaactcta  
29521 gaaactttag caattattag agatatatag cttaagcctt cactactaca gaggccagta  
29581 gttgaattatt aaagcactta aagtatccaa gcatttact tgttcttttg tcatataatc  
29641 ttttaccatca gtttactttc ttttagaaat aaatttttga atttttctaa agttttctat  
29701 ataccattct gtatttatat gacactttat tctacttatt ccttctctat gtgcttctgc  
29761 tagctacaat tcactaaaag aactgatagc gggccgggtg cgttgggtca cacttgcact  
29821 ctaagccctt ttggtggcgc aggcggcgccg atcacaaggt cgggagatca agaccaacct  
29881 ggccaacagg gtgaaccccc gttttacta aaatagcgaa aatlagccgg acgttggtggc  
29941 gtcgcctgt agtccagct actcaggagg ctgaggcagg gcaactcgct caacctggga  
30001 ggtggagcgt cgagtgaact gagatcatgc cagcgtactc cagcgtatgt acagagcgcc  
30061 actccgcttc Naaaaaaaaa aaggaaactga cagcgtatc acgtgcctgt catactaa  
30121 ttctctatc attaaaaaag tcaccaaaat gggcattctc ttltgaaagt ctcagttgaa  
30181 accagctgtt gtttcaacta cagttatgca ggtgtgatt gttttactgg tgtttccag  
30241 taggtatata taatagagac acttaattac tgaattgaca acttaattac ttccagttg  
30301 aaatcagaag aggaattctc cactcttggt gttcctttac agtcatctga gagaagaggt  
30361 aaagaagagg caagtttaga tttaagtct acaagatgaa agttaaaggc cagcagttcca  
30421 tcaaaagacg ctccatatac actgacatat gcacagccgt agttaaaggc cagcagttcca  
30481 ttgttttgct agcactagtt taccatataa ttctagttaa aactttgaac tagtctctat  
30541 aatggaaaat aacttagaaa ttattcttga cagttataaa tgttgggtag tatttttcca  
30601 tatttaagat gatgtgact ttctctttt ttatttgaga cggggtttcg atttttttgc  
30661 ccaggatgga gtgcaatggt gcaatcttgg ctaccacaaa cctccgcctc agtgttcca  
30721 gtagtctccc tgcctcagcc tcccaagtag ctgggattat aggcactgcg cactgcgcc  
30781 agctaatttt tttaaagta aagacggggt ttctccatgt ttgtcagccc ggtcttgaac  
30841 tccgcacttc aggtgatctg ccgcctcggg cctcccaagc tgcctggagt acttgcctga  
30901 gccaccagcg ccggccgatg tgactttctt tggatgaaa attgttcttt tgatatacaa  
30961 aacatatctt aaaaaagtc aaattacagg accatatgaa ccttaattgaa cacttcttat  
31021 aacatctta aactgaaaca gcagaccctg tccattccat taaaagtagg gtttgatta  
31081 ctaataacaa aacttaccagc taacattgta tatacatct catagaactc gttgaactt  
31141 catcttgaaa aataaaaaaa gaaactacaa gtcaagtgtg tgacatgacg ctgtataggc  
31201 atggggaggt tatctataaa attaatcaac atgagttagc aaaaattatt tggactacag  
31261 caaagaatac ctctcgcta ggagaatgca tgacagggcg aggagagaga ggtgtcccc  
31321 ggggaatttt ctatttaac tcgaagtata aaaaacacac cagagagatga tatagtttc

31381 tccctttttg tcatcatatc aagctctctcc tcctttttct tgcatataaag ttccaaataa  
 31441 agacaaaaaa agagactgtg taatgatata caatttaccac tccactttga caatgaataa  
 31501 aacatggcca actttgaaac tgttactcac cccactctca caacattttt gtcatttttg  
 31561 agagtaatat acttttttcc acttttaaaa gcagctttag gtgcgggtga gtcggtcatg  
 31621 cctgtataac cagcactttg ggaggccaaag gtgggcagat cgctcgagct caggagatcg  
 31681 agacactcct gcctaacaat gtgaacccct gtctctacta aaaaatacaa aaactagctg  
 31741 ggcattgttg tacacgcctg taaNcccgag tactcaggag gctgagacaa gagaaatcgt  
 31801 tgaacccagg aggcagaggt tgcagtggag tgagatcatg cccactgact cagcgcctgg  
 31861 tgacagacag agactgtctc tcaaaaaaca aaaaacaaaa caaaaacaaa acaaaaaaaa  
 31921 ggcagcttta ttgagataat tgacatttat tagctataca tatttaaaag atcagatttg  
 31981 attaaatttt gacacaaaaa tgcacctgtg aaatcatcat cacagaaga atagtgaaca  
 32041 tgtctatcat ccccaaaaagt ttctgcttgc cctctgcta atctcccatc  
 32101 tctctctctg cctcactctc tcctccagag gcaaccacot atctacttct tgtaactata  
 32161 ggttagattg cattttctgg agtttttagat aaatggaaac taactataag tactttgtg  
 32221 tgcctggctt ctttcaactca gcatatttgt ttgagattca tccatgttat ttgtttgtat  
 32281 agtagccatc tcttttcaac ttaacttgaac ctggaggctc ccatggatc ttctctgatt  
 32341 tctaccatcc gaataacccct ctgtttagct ccagagttga atgcccactc ttgctgtgat  
 32401 ggttgtgtat tgaagaaagc agataaaaaa taaggttgta atagcttcaa gctttcagta  
 32461 ggtattttct ttctattaca atcttaagct agactacatt ttcaatttct tccgaaagt  
 32521 attaacgaaa aattaacaaa atatttagac aaagcagaca ttaacagagg taaagaaaca  
 32581 taataattga gttacagtca aaatataaac ttgagttgta ttagtttgca taacacccat  
 32641 cctcatatgt tcttcaacat ggaaggtgct gtgttcagtt tcaacagatga cagacttgta  
 32701 aaatagttacc ttatttagat gcctgcatct gaggaaaaaa cactcaataa aaatccagca  
 32761 ggcaactacc ggttagcagt tttagattaa tcataggtta taaaacagag caaagaataa  
 32821 tgacagcagc cctcaccagct ctgaattttt cctcaattat tgtattgtat tcaatttctt  
 32881 atgcagattt agtcaagtaa aacgcctcaa aatttagaag tattataaag atatttatta  
 32941 ctataaattt gaaaactata aataataatt gtgatttagg tcatatacca cagtaactg  
 33001 ttaattttYta aaaaacaaaa atataatttg aaaaatgcat agggccagcg cctgttgctc  
 33061 atgcctgtaa tcccagcaca ttggggaggcc gaggtgggca gattcattga ggtcaggagt  
 33121 ttgagaccag ccttggccaa cgtggtgaaa cctgtctctc actaaaaata caaaaattag  
 33181 ctgagcagtg tgggtgcaagc ctgttaatccc agctactcag gaggttgagg caggagaatc  
 33241 ccttgaaccc aggagggcaga ggttgagtg agccaaagatt gtgccactgt attccagcgt  
 33301 ggggtgacaga gtgagactcc atctcaaaaa aaaaaaaa aaagagaaat agcataagct  
 33361 ggccacagt gctcagcctc ataaccagc cactttggga cctgagagg attgcatatt  
 33421 tgagttcagg agttcgagac cagcctggcc aacatggtga aaccctgtct ctactaaaaa  
 33481 aaatacaaaa attagccagg tatggtgggt ggcacctgta atcccagcta ctcaggaggc  
 33541 tgaggcgaaa gaattctctg ggaggcgag gttagcagta gccaaagatca caccactgca  
 33601 cccagcctg ggcaacagag caagattctg tccccacc caaaaaaaaga cataaaaaga  
 33661 ggccaaagt cgtggtctca gtctgtaatc tcagcacttt gggatgacaa ggcagaggga  
 33721 ctgcttgagg ccaggagtcc gaaaacagcc tgagcatcac agcagagacc cttctctaca  
 33781 aaaaattaaa atattagcca ggcattggtg catcaaggca ggagatcac ctgagccagc  
 33841 gaggtggggg ctgcagtgag ctgtgtttgt gtcaacggtac tccagctgac ccaacacat  
 33901 aactttatct cagaaaaaaa acaaaaacca aagtgagac cttaaaaaat ctgtcttca  
 33961 attgtctgca aagacagtag acaggctttg agtcatatt gaaaacaaat catattata  
 34021 ggttgctaag tgggttagtg aaagaataacc atgaccccc aagctgaatc aagttcagaa  
 34081 ttaactaaag tcttctagcc atgattoata ccgataatac tgagcaggag caaatttgat  
 34141 agctgcctga actggcatg ctggcgagc ctctgtgat acagattttt ctaaggctac  
 34201 ttgtgtcttt tctgttatct gaaaaggaaa atatacatt atatacatt gagatttgg  
 34261 ggtgttaaat agtcatcgaa tatatcaatt cccagttacc tctttaatag gcttctctat  
 34321 ggccctttgc aggtctggat catctgcat atatacatt ctaaacctcc ttggaacca cttcagtgta  
 34381 ttgtcataa atgacctaaa atgttcaaac acaagcaggc ttaagaaaaa caaacacga  
 34441 aaagaaaccc cattagccat ttcagatgtc aacatgcct ctgagcac tataaaaatt  
 34501 aggcactaca gattttctct tttagatttt aaaaacaaag aagcttaaaa taataaaa  
 34561 aattgctacc ctggctgtaa gttagtaatt ttcccttgac ccatgtcata tatattat  
 34621 gagattcttt aatcaagtta gtatccaaat aaacactgca acaaaaaatca tttattat  
 34681 atcactattt ttgaggccca ctaaaaactg aaattaatga attaaatga ggttaccaaa  
 34741 accattaac ttgattacaa actcatcaga tcatctatt tataacgctt cctcttttct  
 34801 atcactggg cctgaaaaat ttattgacc tcatctatt tataacgctt cctcttttct  
 34861 ccatcaactc tattcacaga ctaagaaaaa aaaaatgtgc cacagttaa aaacccagca  
 34921 tgttttcttc catgagaaga aaaaaaaa aaagaaaaaa atgattatg atctactata  
 34981 tatacaactt attttatatc cattatccaa attgaggctt gttattatg atctactatg  
 35041 cccaacatgt ttcttattag cgtgacactg atgtgttagg ctcttaacaa ttgataact  
 35101 ggttactttc aggtccatct atacttctta aatatttca taacaattta catggcat  
 35161 tttaagata aaaaagata cagaggggccc gggcgcggtg gcttaccag

35221	cacttttggga	ggctgaggtg	gctggatcat	gtgggtcagga	gatcgaaacc	atactagcta
35281	acactgttgaa	accccgctct	tactaaaaaw	acaaaaaatt	agccagtgtg	ggtggcaggc
35341	ccctgtagtc	ccagctactc	gggaggctga	ggcaggagaa	agcctttaac	ccggggagga
35401	gagcttgtag	taagccaaga	ttgcgtcact	gcactccaga	ctgggttgaca	ggcgaagcm
35461	ctgtatatca	aaaaaataaa	aaaaaataaa	aaagagagaa	tacagaataa	tcaataaata
35521	ataggaatc	tacccttaaa	gcttttagta	gaattctttt	gaggtacata	aaaaaagaag
35581	acaactttta	aaatcaaaag	aaaaataaa	tcttaggcac	ttagagcttc	cagaattttt
35641	ttttgtagaa	ctgtttacca	ctcaaacgca	aagtcttata	ttaataagtc	tatacatatc
35701	catctatata	aaaagtttaa	aggaataatga	tgtctattct	aattttttac	ctcaaacacc
35761	gcacaaaagt	gtgagccact	ttatggctta	tgcacattga	caactctttg	ctctagcctt
35821	caactctccc	ttacaactct	caactctcta	ctcattatgc	ctcagtcaca	ctagctcatc
35881	ttctgtttct	tgaacataac	atacttttcc	ccaattctag	gcctttcacag	cttctgtttt
35941	atcattcttc	aagttaagctc	tcaatctgtg	tgtctcagat	caaaggtccct	ccatttatcga
36001	ggcttcttct	gatgtgcctta	gtagtaaccc	ctccccacc	aactcatttg	ctctatagc
36061	tgttttactg	gatcaagact	gttcacagca	tgaattatc	tacttttacc	ggtttattatc
36121	tgtctccgtg	atataaacaa	taaaaacggg	gagggcacgg	tgctcacgcg	ctctaatccc
36181	agcacttttg	ggggccggag	tggtgtggat	acctgaagtc	aggaattcaa	gaccagcctg
36241	accacacatg	tgaacccccg	tctctcccaa	aaacagaaaa	attaaccagg	catggtgtgtg
36301	gggtgcgtga	atctcagcta	catgagacgc	tgaggcagga	gaatcgttgt	aaacccggag
36361	agccacttga	cagtgagccg	agatcatgcc	attgcactcc	agcctgagca	acaagaggtg
36421	aactgtcttc	aaaaaataaa	aaaaaataaa	aagtgtctcc	tgtgtcttat	ataatagact
36481	ggagatatatt	gttcaatatt	ttttgtacaa	atgaataaat	aaatgaattt	tagagatatg
36541	ctgtttacgt	ctaaaggtgt	ttataggcca	tgaacagtga	ctcacgcgtg	taatcacagc
36601	ttttgtgaag	gccaaagttg	gaggtatagct	tgaagccagag	tgtgatgtga	gaccgacagg
36661	aaacatggca	gacctatctc	tattttaaag	aatggatgga	tgaattgaata	aaataaatag
36721	ttttatattt	agggcgtaca	gaactttcta	taagctactg	atttcccgca	aaaaactatc
36781	tcatggctat	acactaaaaa	ctccagcttg	tttttctcag	gagtgagaag	agggatcagg
36841	taagaaagcc	taagatgatt	tgcacaaaaa	cagctactac	tgtatgtgtg	ctttataggt
36901	tcaccatact	ctaagaagtg	tgtattttga	agggaccata	aaactcacct	tactgaaaaa
36961	agtgaggttc	agtaacattaa	gatcatacag	ctactatctg	acctaagatt	cttttttttt
37021	tctcagctat	gaaaaacccta	tacaacccta	agctgaaatt	tatatataaa	cacagaaaaa
37081	aaataaccaat	tctatgaaga	gttaactcatg	aRgagtttat	caacttctcca	tgaacaaggt
37141	tataaggaaa	ggattcttaag	taaatggctc	cttggtagtc	atgtatgttc	aaacataaag
37201	ctcaatactt	atttagcata	caagagtcac	tttgcgtccc	ttattttgggc	attcccaagt
37261	taataaccctt	tcaaatgagc	actaRtaaga	ttctagtata	tcacaaagct	gagtgaaagg
37321	catctgaaga	cgaaagttcc	tccatagtaa	ttaatataaa	aaacccccat	tttttttttt
37381	ttgagacgaa	gtcttgctct	tggtgcccaa	gctggagtg	aatggcgtga	tctcagctca
37441	tagcaacctc	cgctcccccg	gttaaaacca	ttctcctgcc	tcagccctccc	tcccaagttg
37501	ctgggatatac	agggcagatga	caccacgtcc	agctaatttt	tgtattttta	ttagaagatg
37561	ggttttaccga	tgttgccagc	gctggctctg	atctcctgac	ctcaggcgat	ccaccgcctc
37621	cgccctccca	aagtgtctggg	attacagggc	tgaagccagc	caactgcgca	aaacccccaa
37681	tttttaagtg	ctgagaatta	aagtgtctga	ggaaactgag	taccaggagtg	caatgggata
37741	ggggagYgtt	ttctgaaaaa	tacaataatc	ctctgaaagt	atcagaagca	caataaagct
37801	tgttcaacaa	gataaagtag	taactcatct	gtgaaacata	tatttttata	cgaaaaactg
37861	gcacaaagct	agactgacta	aaaggataag	ctctgcacaa	catctacatt	gtgagatata
37921	aagaaaaagc	taactagctg	tagctagtgt	gtgaaatcat	acaaacaaga	taacatagta
37981	tttatcatgg	ggctagttta	aagtgtttac	gggttaattag	gggttaattag	ttgaagcata
38041	ttctttgggt	atagtaYgtg	tgtaaacttt	cgtttttcta	agagataggg	tttcagttat
38101	tgcgccagcg	tagtctcaaa	ctcctggcct	caaatgatcc	tcccacctca	accttcccaa
38161	agtgctggga	ttacagggcat	gagccaatgc	acttgccctg	tatgtagtat	ttcttagtat
38221	actatttcag	agggaaaaat	acatacacca	agccctttat	tcaagagttt	attacttgcc
38281	tggtgtgtgt	ggctcatgcc	tgtaatccca	gcactttlag	agggccaggt	gggtggatga
38341	ctcagaggtg	ggagttggag	actagcctga	ccaaacagga	gaacacactg	ctctaccaaa
38401	aaatacaaaa	tagtcaggcg	tggtggcaca	tgccctgaat	ccagactact	caggagaggt
38461	agggcaggaga	atcacttgaa	ccggggaggg	agaggttgcc	gtgagctgag	atcgtgcatc
38521	tgcaccccca	ctctgggcaac	aaagacaaaa	ctccgtttca	aaaaaaaata	aagatttatc
38581	acttttcata	ttaggaaaaa	aaaagcattt	ggttgatata	agattttcac	taaaaaaaac
38641	tttttcatcc	ccaaattttgt	atgactaata	agtcgtatac	acagctacta	cttcacatgt
38701	agagaaattg	cttagacttg	tcttagaata	ttacgtctta	gaactctcac	tgtatctctc
38761	caactcttat	acctaaccac	tagtattttc	catcttactc	catcactatc	taattattca
38821	gtgagaatcc	cctcgtttct	gaagaataac	agacctctMR	gaagcaaatc	agaaattctc
38881	ctgcacaaat	ataaatccaa	taacctgttc	agtccttgga	ataacaaatt	tttttaactt
38941	tattttatata	attatgcagt	ccattttcat	ttcttctaaa	acagagaaat	cagtgatgtg
39001	caaacaggag	atcactgcga	cctctaactc	ctctctccca	gtcgctttgt	aaatgtgtga



39061 ggtgtttctta gtttaagacaa tgggggttagg atgggggacgc taactgggatt tatagtgatt  
39121 gccacttatg ctaaacattgt tacaatgtctg aagatgtctc taacaaactcc aagtcctccc  
39181 ccaggtgccaa tggcacataa tagttactctg attaatgttg aattactcaa tgcattccaa  
39241 acctctatgc tgaatctatc caactctaga aaaaaacaa acaaaaagatg ataactctga  
39301 gagggtgtccc agttttgggtt ttgtttctatg ttggtgcatt acagacaaaa ccgaactgtc  
39361 gcccgggtgt agtcctatct gcctgtaatc ccatcacttt gaggaggttgc agtgagccaa  
39421 gatcgtgcca ctgcatttca gcctgggtgga cagatgtgaga ctccacttcca aaaaaaaaa  
39481 aaaaagaaaa aaaaaaatag cataggtctgg gcacagcac tcaacRcctat atatccagaa  
39541 ctttgggaag ctgaggaggga tggatcattt gaggtcagga gtttaagcac agcctgggcca  
39601 acatgtgtgaa accctgttttc tactaaaaat acaaaaaaat ggttagggcca ttgtgggttg  
39661 ccctctgtaat ccagctacag ttggatgggt agggcagaaga atcacctaaa cctggggagc  
39721 agagggttcca agtgaccaa agttgactcc agcctgggca acagagcaaa gcttccattc  
39781 aaaaaaaaaa aaaaaaaaaa gaaaagttaa gaaaagaaaa gaaaaaaaaa agaaaaaaca  
39841 aggaacaactg ctttggagac cttatttggc caagatacta aactccaaa atagatattg  
39901 cagtcctgaac agttctactc tattttgaat acaaaaat at caccatata gacatgtg  
39961 taaacatcat tccctagaaa tcagataaaa ctctgcttt ggttaggaga ggtgaagatt  
40021 aaaaaggacc aaagctgtat ttcattcaaa ataacacaa ctgcacatgc caaattttg  
40081 cctctagatt ttgttttttc ctaccaaatg ggaatggaga ggtgacact accaataagg  
40141 aagatgtatt acccccaagg agaaatgcac caaaaaaag gcaactgag tcaagtgaag  
40201 gcctccatgt tatagRgtaa atttccagta tatgttctcc atgtatagg gctcaatttt  
40261 cctttttcag gagggtgacta Yacttctcca ctggcctcag aagtggtttt ttgttataca  
40321 attggaaacac ctattatttg ctgaacaaac actgggtggac gccttatgtg tcaacttct  
40381 aggcagtaag taaggcacca aagacggaaa tataaatata taatccctgc ccttaaaagc  
40441 attatagta caatgatacag attatgagtg aacagataat taacacaaaaa ggtgaactag  
40501 gctatgataa acattgagag tagagaagag ggaactcag tctagccagg gaaggtccag  
40561 gacaaacacc ttgaatagtg aagagtggag tgagaatgga aactcaataga atttaagtac  
40621 aagaatgaata agggtagaac ttcttaggaa aaagaatata cctggggaag ctcaagagag  
40681 aeatgtgtcat attatgggaa ctacaagtag ttctctatgg caaatattta tatgaagatt  
40741 ttgaatgttt ttcttctctg gctggggcag gtgtctcatg cctgtaatcc cagcaacttg  
40801 ggaggtggcc cagggtggac tgcttgaggt caggagtcca agccagcact acccaacagc  
40861 gtgaaaccca ttctactaaa aaagacaaaa acgagctggg catggtggcg gggcctctga  
40921 attccccagtc ctcagaagcc tgaggcagga gaactgtctg aactgggaag ctgaggtgc  
40981 cagtgagctg agattgccct actgcactcc agcctggggg acaaaagcaag atttctctc  
41041 agaaaaaaaag tttttttctt cttaacaacc accctggcta ggtgcagtg gctacacactg  
41101 taattccaac actctgggag gccaaagcag gcagatcact tacaaaaaa ttgagctcag  
41161 cagctgtggc acaaggtga aaccctgtca ctacaaaaa gctaccagc ctgggtgaca  
41221 ggtggtgtgc acctgtagtc ccaactcctt agaaagctgg ggtgaaagga tcaacttgagc  
41281 tcaggagttc aaggctgcag cgagtctgta ttgtgccact gcactccagc ctgggtgaca  
41341 gagcaagact gtctcgaaaa caaacaacaa cacatataat agtatttatt ttatagtgaa  
41401 aaaaaacaga atggctaaaa ctgaactatt tatttttggg ttctagttatt ccSatcagga  
41461 gaggttgacg tagccaagaa aagtgagtaa ggggtcaagc atgccatagc agcttatgctc  
41521 taccctaaa gaccatggaa gccatagaaa gatttttaatt gtgcacact gtgctgatgt  
41581 gtatatacat gatgagaaag acaaacagac atggatgaat ttgtttgttc atgccagcaa  
41641 gacaaatgtc aatgagatgg ttgcaggctt ggctcagta aatcaacgaa taattgtctga  
41701 cagcctgaac taagtacag agaggggaat atacttgaga agtgttaagg aatataagct  
41761 gtctcacag aataaataga tgaagggtct gagggaacaa agaactagaa ataatgtctga  
41821 ccagagagaa catatgggca gtatgtgtca catgtgaaat ctaaatagcc tttagccag  
41881 taagtccaca cctaagaatt taaacaaaa atcatgttca acaagatgtt atgtatgtga  
41941 tcttgaatat atttggttgt aaaactaaaa accaccaatg cttaaaaaaa ctgacagata  
42001 accatgccac ggaactctat atagattaat aaaaaaaaat gaagttagc tatatgtact  
42061 acaaaaaatt gtcaaatgtg ttaagtaaaa Waaaaagcaa gctcaaggaa gttatgggca  
42121 tcaaaattcc attcatgtaa acaacagatt ctaagtttgt atgtatacat atgtatggca  
42181 tatacataaa tatctaaatc cataataaaa attttggaatt atagctcccc ctccccctc  
42201 cccctctctg gcacggtctc cctctccctc tccctctcca cggctcccc ctgagctctc  
42301 gccgaggtgc gactgtgctg ccacctctc ggctcactgc aaactccatg cctgattctc  
42361 cctcctcagc ctgccgagtg cctgggattg caggctgcgc ccgccacgcc ctgaggtttc  
42421 tctgattttt ttggtggaga ttgggttttg ctgtgttggc cgggtgtgct ccagctctc  
42481 aacagtgtgt gatctgccag cctcgccctc ccRaggtgcc ggtatgtcag tccagagct  
42541 actcaactcag tgcctcaatgt tgccccagct ggagtgcagt ggcgtgatct cggctcgtc  
42601 caacctccac ctccccagcc cctgccttgg cctccaaaag tccgagact gcagcctctga  
42661 gccggccgca accctgtctg ggaagtgaag agcgtctctg ccaggccgca catcgtctctg  
42721 gatgtgagca gccctctctg ccggccgccc agtctgggaa gtagaggagca cctctctccg  
42781 gccgcctcat cgtctaggaa gtgaggagcg tctctgccgc gcgcctcact gtctgagatg  
42841 tgggtgagcg cctctgcccc gcgcccgtc tgggattgga ggaagcactc tggccagcg

42901 gggccctgtc tgggaggtga ggagcgtctc tgccYggcca ccccgctctga gaaggggaaga  
42961 gcccctccag ccggcagccg tcccRctcgg gaagtgaagg cagcRctccg ccggcgaaga  
43021 ccccgctccg gaggaggtg ccggggcagcc cggcctgtcc cagccgcccc gctcgggagg  
43081 gaggYggggg ccagccctg ccggggcagcc cggcctgtcc gggaggtggg gggcgctctg  
43141 cctcggccgg ccgcccctcc gggaggtggg gggcgctctc gggcgctctg cccgctctgg  
43201 aggtgagggg cccctctggc caagggccac cccgctctgg agtgtatccc aacagctcat  
43261 tgagagcagg ccatgatgac Ratggcggtt ttgtcgaaat gaagagggggg aaatgtgggg  
43321 aaaagataga taatcagat ttgtgtctgt tctgtgtaga gggaggtaga cgtagagagc  
43381 tccattttgt tctgtactaa gaaaaattct tctgctttgt gatgtcttta atcataacct  
43441 tacccttaac cccatgtctc ctgaaacatg tctgtctatc actcagggtt aaatggatta  
43501 agggcggtgc aaagtgtgct ttgttaacaa gatgttgtaa ggcagcatgc ttgttaagag  
43561 tcatccacc cccatattct caagtaacca gggacacaaa cctcgcgaga gggcgaggg  
43621 tccctgtgct agggaaacca gagaccctgt ttccattgtt tatctctgta ccttcccccc  
43681 actatgtccc tatgacctg ccaaatcccc ctctYgaga aacacccaag aagcataaat  
43741 aaatactaaa ataaataaat aaataataaa aatatttttag gaaaaaaaaa tttggaaat  
43801 tataatacac tttagtaaca gcagttaccc tctgggaagt aagttaggga actcaggcca  
43861 ggactRtata tacttgttat atatacttat ggtattcttt actataatatt attatttgta  
43921 taactttttt tttatttga gacagagctc tctcgtctgc cccagcgctgg agtccggtg  
43981 tggcctcttt agctcaactg aagctctgcc tcccaggttc agccactatt cctcgctcag  
44041 cctcccgag agctgggaat acaggttacc gccaccatgc cactgaattt ttttgtttgt  
44101 tttagtagaga cgggttttca ctgtgttagc caggatgttc tcaattctcc gacctctgta  
44161 tcccaccacc tcatctccc aaagtgtcgg gattacaggt gtgacacca acccccagc  
44221 tttttgtata attKttttta aaaaaataa tataatagta agaaaaaatg actccccag  
44281 tatctcagata acaaatagg aaaaatacac taccacttac tgaattacag aatagcagg  
44341 tagattaaat ggaataattat gtattttagg ttggaattga ttaagtatga ctttagggaa  
44401 actatagata acttgcctct gtatagacgta agatataagg accctacagc aagcagaga  
44461 tactggggtg gaaacacaga ttgaacagtc atcagcatat gaggtygcaaa taatatacat  
44521 taactttttt aaagtgtcgt tgttttttgg tttaaagtcac aagatgataa aagaagaaat  
44581 atatatataag agaattRtgt gtataaaaaa gtgttaagag actttttgca gagagacct  
44641 ccaataaact gtttaggtga ggtcattttg tagtgaattg aaagatgaaa gaagaagtga  
44701 gaacacagcc aaaaaaact cctgaaacaa gcttgaaaaa agtgaagaaa attagtgtgt  
44761 agctagacag aacactgcga gttagttact agggaaaagc ctatctttca aaatagata  
44821 aaggccagcg agcctcatgc ctgtaatccc agcacttttg gaggccaggg Yggcgagatc  
44881 acagagctcag gagatcaaga ccatcctggc taacatgggt aaaccccatc tctactaaa  
44941 atacaaaaaa attcaccggg catggtggca ggcgctgtg gtccccagta ctcggggggc  
45001 tgacRtaaga tgatgcctg agcccaagag ttcaaggttg caatgaattt agagattgtc  
45061 agggcactac aatctaaaaa agaataaaca aacaaatgac aaagaaaaaa agcacattca  
45121 aatgctaaca ggaagagcc gttagaagag gagggatgaa gaactgttga gaagattaac  
45181 tagtagtagta ggttcataga ggaagagggg agggattggg atccagaata catatggcaa  
45241 gatcaggtgt agatccaggt gcagtggtcc acgctgtgta tMccagcact ttgggtggca  
45301 gaagcaggag gattgtctga gcccaaaagt tcaagaccag cttgagttaac ataggvata  
45361 ccccacttca caaaaaata aaaaattagc caggcatgtt ggtgcacacc ttgctctca  
45421 gccactcgag aggtgaggt gggaggaatg cttgagccca gggagtcag gcttcaagta  
45481 gccatcatgc tgccactgca ctccagcctg ggcaacagaa caagaccctg tctcaaaaaa  
45541 acaataaaaa aataaaaaaa aaaaagatta gcttagaca gaagaaagca ttctgtgcaa  
45601 aaaaacaaagt aaaaagtcag gaaaaggagc aagtgacgga tgaagaaatg tgaacaattg  
45661 ctccaatgat tctttatatt caaatagat gggagtgca cctctgaaga atgatgcaga  
45721 taagagaggt aggaattaaa gtttgtgtgg agtaccagg gacttagact gaatatggag  
45781 atataaaatt gcatgattgc ataatgtct tccccaaaag ctacagcacc atgatgcaga  
45841 agcaaaagat ggttgtttag tgcattttta gagggtaaa ttgaaaagtt tggccggggt  
45901 agagagggcca atgataggca ctgaaagctc tcaagtattt ttccaaaact acatagttc  
45961 aaggaaattt ttcaatttca tcttgagat ttattttaca gaaaaaaaat aaactaggtc  
46021 tcttataaaa tagttcctgt gccctgagtg tagctgttct gtgtatgagt atttttaagc  
46081 aaattttttt ttggccctat ggattacttc gtatcagggc catctcaaal ctgatataga  
46141 aatgctagt ttgtacttaa tctagggcaa aagaagtttg tglattttat taaactcaa  
46201 taaactatta aatacttgta atagagtaca tglagtact ctaagtacat gggagaaagt  
46261 agtcaagact ataggaaaaa actagaagaa actacagagt aaattaaaa ttgatgttaa  
46321 gctctgggg ttgcagattt ttaaaaattt actggcagt ttagaattgc atcccaagca  
46381 agtaactgtt ttaaaaatgta ctttttccc cttttaaacc aaggttatgc ccccaagca  
46441 agagtgctgt gctcactgt atcctcaaac tcttgggctc agtcaaacct cctactttaa  
46501 cctcccaagt agctgggaat ataggtacga gccaccatgc ccagccactg cttactttta  
46561 cataatcaca acagcaagag atttatcgtg tccacaaaat ttaaaattgt tcttaataa  
46621 tttatcatc agggaaaaa tgagagaggt acctgtctt ttgactgtcc ttgtcagca  
46681 attgcatcat atttaatttt tccctcagaa tccacctgaa tggccagcgc atttcgactt

46741	ttttttcttttc	gtcccatatc	cagtggtatc	tgggcccacat	ggatctcttgg	aaaagccact
46801	ccatctcccaa	aatccctacac	attagaagta	gttaataaaa	atagtttaatc	atctatttgac
46861	atccagataga	tattcattcca	ttcaaaacagt	aatgtgtgag	tctcttagtat	gctagatgct
46921	agagatcacag	cagagttccc	tgccctcttgg	ggagtttaca	ttcctagagg	caangataata
46981	aaaaaacatc	tctttgctct	gtgataagta	ccatgaaggt	aatataaaag	aaaaagactg
47041	ggaaaagggg	ctacttttga	taagggtgggt	ggggcagcc	tttgtgaagg	agtaacatgt
47101	gagctaaagc	ctgaatgatt	agaaaagcgt	ggcgggaaga	cacacacaca	cacacacaca
47161	cacacacaga	cacacacMcC	acacacacac	acacacaaag	catctttctag	catctctcaa
47221	tagccaatgc	aaaggccctg	ctgcagaagt	gagcttggtt	tatttttaggt	aatgagaatt
47281	cagttggggct	aacagcttaa	taaggacagc	agagcttgag	gcacatgtag	catgagttaa
47341	aaaacacaga	ctcaaaaagg	acttcatctg	tacagcaga	aaatgtggat	ttcatatagt
47401	caggacgaaa	ctaagtaaaa	tggttaggac	aaaggcaaat	atatgtatag	atattttagc
47461	atataaatac	atattttagt	gagggaacagg	taaaagtata	cttggggcag	aaataataaa
47521	gattttttct	tggaatcaga	ttttattaac	atattattta	ttcattgtta	tgtagaagc
47581	aaatgaatac	aatagggaag	atccagaaat	gagtttacta	catcttctact	agtgattata
47641	cataattatt	taggagagag	aactgagact	gctaaagtga	aacaaagtaa	ataaagtttg
47701	ctgataaagt	gtagattttt	aaaaaggttt	atcaaaataa	catgaggata	aagtataatt
47761	tttaagtagt	taaaaaaaat	tcagtagtat	tttgttagta	catctttaa	atgatKcaca
47821	catgtattat	ctcatttta	ccttgtta	ctttaacat	attttaaatg	gaacatctca
47881	ggccttaaaa	ataaaagcgg	tggttctgag	atctctatg	tattatagtg	agacactgaa
47941	ccagatcttt	ttgagactcc	gaaaagcttt	gtattttcta	ctactatgaa	atacttttct
48001	tttagctcaa	aacagccaaa	tgagtactat	taagaaataa	tagaaaaaac	aaatataaat
48061	taattcattt	gaggggggtt	gatggaaaaa	aaagaaaaat	ttcactttta	caagaccttc
48121	taagataata	ctcYRttaga	tctaagatat	atgagttcct	aaagtaacaa	atattttttg
48181	ctaaagagat	gtttgaaagt	tttgaaaacc	aaagagaata	agctgagaga	catttttcag
48241	taaatagaga	agagtttaag	catctctaag	ataagKgtct	ccgtaccaca	agccatcaat
48301	ctgacacatg	cttgattttt	ctagagtttt	tcaactataa	taagtaatttc	agaaagttag
48361	ctgtagcatt	atgtggctat	taagcttgcc	agaaaaatcY	agctttttat	ttgaaaaatt
48421	gaacaaaagt	cctctctcca	cagaacattc	actcccctca	agcagactcc	atagcatttc
48481	cagattagct	agccatttca	acttgatttt	caaataaagt	gttctgtttc	accaggattc
48541	atgaaaccac	tgatagccaa	gtaagaatct	aaaaatctaa	tagcaaccga	tactctgtat
48601	gctgaaatgc	atggatgaga	tttacctctc	atgtaaatgt	atgcctctga	ccctcatagt
48661	gcaatgtact	gacagaccaa	acaataat	taatgctcta	acagagacac	tcaactgttc
48721	tattatagta	tgccaatcac	tacttttttg	ccccaggtaa	gactcctgtc	taactcttca
48781	tcttctctgt	cctcccaYaa	atWacactat	cagagagaac	aaaagcaaca	aaaaatccct
48841	aaaaatttga	atatgactgg	ccttcacaca	ccagcattca	taaggaccag	aggaagactg
48901	acaaaaaagg	aacaaaaaga	cctacatgtg	ggtttttttg	aatacatagg	acaagatagt
48961	atctagtgtg	gtgttctctc	tataaactct	aataaaaggt	gtggggcagg	cgcggtgctc
49021	cacacctgtc	atccagcac	tttaggagcg	cgaggcaggt	ggatcacctg	aaagtcaggag
49081	ttcaagacca	gcctagcaaa	catggtgaaa	ccctgtctgt	actaaaaata	caaaaaattt
49141	gctgtgRtgt	gtggcacgtg	cctgtaatcc	tggtctactg	ggagcctgag	caggaggaat
49201	cgcttgtaacc	caggagcgga	aggtcgcagt	gagccaagat	cgcacactgt	cgactcacgc
49261	tggaagacgg	agcgagattc	catcccaaaa	ataaaaaaaa	taaaaatttaa	tttaattagt
49321	taaaagatcc	tgagtaatta	gaataggtgc	tgccaaggaa	gcttgttctt	tttattttta
49381	ttttttattt	tttttgagat	ggagttttgc	gttgtgtgcc	caggctggag	tgcaatgggt
49441	cgactcttggc	tcaccacaac	ctctgcctcc	caggttcaag	caattctccc	ctctcagcct
49501	ccggagttagc	tgggattaca	ggcatgcgcc	accacacctg	gctaattttg	tatttttagt
49561	agagacaggg	ttgtccatg	ttggtcaggc	tggtctcgaa	ctcctggcct	caggtagctc
49621	gcccgccttgc	gctcccaaaa	gtgtcgggat	tacaggtgtg	agccactgcg	cccagccgga
49681	agctttgttat	ttggcaacaa	gctagagata	gaaagagcta	ttccagatgg	tctctctccc
49741	ctcattgcct	tcaaatcaac	caaattttgt	cttatcccg	caaaaataa	tttaacttat
49801	atcatagata	tattttaaca	ttctcaaat	aaaggtaaga	aaacaggtga	aggaagaaag
49861	atcagcaaat	gaagtgtctac	tggaattttc	cctgaaggaa	agtgttctct	tcagggttat
49921	gtggtattat	cggagggggg	gaggagggag	aggaagccaa	aatagaaaWc	catcttcatc
49981	ttctttatct	tttactttcc	catttatttc	cgtaattctc	gaatttataa	aatatagttt
50041	ttcttgttaa	ttttaaatg	tttgatatag	gaatagacaa	atggtacatg	tctcttagtg
50101	ccacaagaga	gtgatagata	ataagcaaaa	agaaaaatga	acatcttctg	agggaaatat
50161	acatcttgag	agcacataac	cagaattttc	atcacactga	catctcaatt	gtatcgtatc
50221	aaactttaaa	tttatgtttat	agtggtctat	aaataatttc	agcacaattt	aaacaattcta
50281	aacttaagat	ctatatgtac	ctctaataac	cgaggtatcc	aggttttccg	gtatccgtac
50341	ggggggaggtt	ctcttcggga	ggagaccagt	gaggtctgcc	gtgactctcg	ggatcttgcc
50401	ttcttctcag	ctccaagctg	gtctcgagat	agctgagtag	gtgacagtga	aaagctaata
50461	gcaaaagata	taaaagttag	aaattttaaa	aactcttatg	tttaatttga	ccactgcaac
50521	ttggccacag	agacaatttt	tctacagatg	ctttactctt	tcagaaaaag	agcagatttt

50581 tcaatgatt atggtagtac tcaactactag ccattttata atatgtggaa gacagaagta  
50641 tcaaaagata ccagataaaa acatctgact aggcacaacag caaatgagt cagaataaa  
50701 aaaaataacag atatttcaca taacagacaaa agctctaaag gtttatgtctt ccaatatttcc  
50761 tttttttttt gagaagaagt ctcaactgtt tgcccaggct ggaatagact ggtgtaactt  
50821 ccaactcactg caactcttgc ctccctgggt caagtgtatt tctccctcca tgcctccaag  
50881 tagctgggag tacaagcaat gcgcccaca caggtcaat ttttgcaatt ttgtagaga  
50941 tggggtttca ccaatgttgc agggctgtgt ttgaacttga cctcaagtga tccgtgtacc  
51001 ttggctccac aaagtgtcgg aattacagac gtgtgccacc agactctgcc acgtctccrt  
51061 attYccatatt ttcttctctt tttttttttt gacgcagagt ctctctgtgt cacttaggct  
51121 agagtgcagt gccaccatct cagctcaact cactctgggt cactgtgatt ttcacttat  
51181 tcttgcctca ctcccccag tagctggagt tacagatgtg cagcaccacg cccactaat  
51241 ttttatttt ttatagaga cagggtttcg ccatgttggc caggctgtgt tcgaactccc  
51301 gacgcgagt gatgtgcccg cccggcctct ccaaatgtgt aggatataag gtgtgaacca  
51361 ccaagccag cctgtttcca tattcttat aacaaacaag taacattccc taactattca  
51421 gggataaaact gtccagata tacttgcata tcttatact cttttttttt ttgagacaga  
51481 gtctcgttct gtgcccaag cttgagtcca gtgtgtcaac ctgcgtcac aagaacctct  
51541 gtccgtctgga ttctgtgtat tctcatgctt cagctcccg agtagctggg attacaggtg  
51601 caagccaca cctccctatt ttttatttt ttatagaga ctgtgttttg cactgtgtgc  
51661 ccaatgttgt caggctgtgt tcgaactcct gacctcaagt gatccataag cctctactct  
51721 ccaagtagct gggattactg gcatgagcca ccaagctgtg ctactcctgt tttttaatct  
51781 ttttaccatt atgtcctcta ctttaaaact ctctcagtta aatacaaaaag taactgattt  
51841 aaaaactgaat gtaaaaaaga ggtggataaa ataaacaaa ttgcatattt aaagtctcaa  
51901 gggatctatca aatgaaattt aaaacttaga ggccagacca gccatggttg ctcacactgt  
51961 taactccagc actttgggag gcgagggtgg gtggatcact tggagtccag actcaaaag  
52021 cggctgtact aacatgttga aacccgtctt ctactaaaaa tacaaaaatt gccagcgctg  
52081 ctgggtgtgc cctgtaatcc cactcgggag gctgagccag gagaatcact ttatccgtg  
52141 aggtggaggg tacagtggc caagatttga ccaatgcact ccagcctgag caacaacagc  
52201 aaaaactccat tcaaaacaaa caaaacaaa ccaaaaaacc cttagagccc aagaatcca  
52261 agcagaagcc ctgaagatgc agcttatctt ttatgcaagg taacctttct actaaatggt  
52321 tcttaacaaa ttttaagtat taagaattta cagtgtactt ccaaatgatt ccaaaaggat  
52381 cctgtaggat gaatatggag aaaaataacct tagttccttt gtgattcttg gatagaaag  
52441 atacctgaat agcagagctt attaatattt caaaaatatt tatattttg tctcaagcac  
52501 tatgtcagag atgagactat aaatggagaa agacagetta gtcttgcaga cagagaccaa  
52561 aaaaaaaaaa atcacacaaa tctataatta cacattatgt taagtattga ttgaaaaata  
52621 tagatgctat gacagtctac atgagagggg tctaatccag tctggggagg taagttagt  
52681 ctaggctcag aattcagatc cgacctatgc cacttactaa cctggataag ctacttacct  
52741 ctctgaatct gttttctcat ctgtaataat agttcataag agtatctact tcataaggtt  
52801 ctttaagcta ttaaatgaga gaataccatg aaccagtgg tatagtactt agacactttc  
52861 aatacagggtt ctgtatgttt gattatttca gggctgttcc atgactcctc cagtttgggt  
52921 taactgcctg tctctatgct ccccaaaaac gctctacata ccaactac acacttacct  
52981 tactgtattg caatcatcta ctcaattaac tatctgttt tcatgagact gtgagcagct  
53041 ctccaggaaat cagccttat gcttctttca tccctaagtc tccaagactt ctgactcaea  
53101 ctctcagactc cttaaatgtt gtaggagaaa taaatacact cttcccccag actccaccat  
53161 tttatttcaac tacagccttc ttatcttttc ttaaaacagc cttctcagga tctttacatt  
53221 ggctatttctt tctgcatcta aatcaatcta atcatctatc tatctatcta tctatcttct  
53281 tatcWabtg aatcaatcaa tcaaatattt tctctattgc taataacaaa tccaatttat  
53341 tctctatcta tctatctatt gcaatcaact acaatttttc cggctgtggc aggtgtgcta  
53401 caactgttaat tccaacactt tgggagggcca aggcaggtga atactgtgag gccaggtt  
53461 cgagaccagc catgccaaat tagtgaacc ctgtctctat taaaaacaca aaaaatagcc  
53521 ggcgtgtgtg gcgagcgctt gtaatccag gtaatccag ctgagcgag aaaaactgtt gcaactgtgga  
53581 ggtggagtgt gtatgagct gagattgtct cactgcact cagcctgggt gagagtgaga  
53641 ctctgcccca aaaaaaaa aaaaaaaa gaaacaaatt tctgtatgct ttatagaaa  
53701 ttcttccca gatctccaga tcttttcttt ttgtattgt cttctgtggg cagggtctatc  
53761 ccaagggcag tatgcttga gtacgcccag atcttctcaac agctgtgcat caggctctatc  
53821 tgtctctgtc cagtttacct tctctagagc ttttttctga ccaacttacc caaaatagcc  
53881 acttccccc aactctgag tcaactctta acccggtctt tatttttact ctatgactat  
53941 taactgaaa ttgcattatg attttactca ctctgccta ttgaattga atgtcaattg  
54001 aattctgtcc cactagaatt ccaattctgt ccaattagag tcttaagctc attgtgtag  
54061 agataaaaa taaaacaaac tcaatttctc ctattattat actcaacaa atgcataaac  
54121 caacacttct gatccagat gtgtgtaggt ttttccccac atatacaagc agctatcaac  
54181 tgtgtctact ggaacacagt caagtgtcct caaattcaat tcaattttga tcttatctac  
54241 ctggagatac gtgcagatgt acaggttga aggcctcagt ccaaaagact gccttactt  
54301 cagatgccaa cagtaggtcc aaggccactg gccaaactac tcaaaacagc ggttccatgt  
54361 acctcatctt caggttcaac taatttgcta gggcagctca cagaactcag ggaaatgtgt

54421 tttactgattt aaagcatattt acaagaataa ctgaggaaaca ccagatgggta ggtatggaaa  
54481 gctctggaggt cctcaagactt ggcattttaa agatcgcatc ttgcccactg aatttttgatc  
54541 acgcaataacc catgacaggY catgatgatgt ggtggctcac gccagctaatg ccgacactgtg  
54601 gaagcgtcga taggggtaga ctgactcagt taagcccagtc tcaagaccagc cgtggccaac  
54661 atggcacaacc cccactctcta ctaaaaaaaa caaaaatttaa cctggcatca tggcagacgt  
54721 ctgagctccc agctactctta gggcgctRag ctaggagagt ctactgagcc ttggaggggt  
54781 gaggStgcag tgggctgtga ttggggccact gcactccagc ctgggcaga gtgagaccct  
54841 atatcaaaaa ataaaaataa taaaaaatac ccatgacatc tctagaattt ttctctgtgt  
54901 gtccattttca aattccacctt ttccacaaaa gtctcccctc ataatctttac tagtctatga  
54961 cacagtttct cacacctggg gagcactcaa atgtttgtgt cctgtaacct tactaataca  
55021 aactctttgac tgggcatggt agctcatgcc tgcataccca gcactttggg aggccttaggt  
55081 gagaagattg atggaggcca aSagttggag accactctgg ccaacgcagt aacacctatg  
55141 ttctagtttt tttttaaaaa aatttcagat gccgggcatg gtgctcatg cctgtatacc  
55201 tagcacttcg ggaggccgag acaggaggac tgcttgaggc caagaattca agaccaacct  
55261 ggccaacaca gcgagaccgc atctcaaaaa ataaagttaa aaaaaaaggc aaaaaattt  
55321 tagaaacgag ggtggggaga aaactctcca gccacataa cttattccat ggtctctttt  
55381 ttgcactaag gttagtatta cacaatagca ctttttgtgt tctgtatgtt tttttatgtc  
55441 gtatactctg caccctatc ttccctgtgt ccaaagacag attttttctg gtgtacacaa  
55501 cccccaggag agagcaaggc accaacagta attcccaaaa taattttttt attaaaaatt  
55561 ttggagctac actgtcatca gccagaaaaa aaaaactgct tgaggagatt gttgaggtgt  
55621 aaaaagactc aagctgatgt cctRaaagga ttactttatc tataaccttc ctgaacacta  
55681 gctttttctc gtgagttctg catgKaaagt ttctccactt aaaaaggggc agcaaggtgt  
55741 ggaactagta ggaagtgcgt aactctctgc actctccacc Tgaaaagcat cttttttcta  
55801 aagggcggaga gcactttaca gacggttact aaggtaacgc Raaacagcaa aactgaggt  
55861 gtctcatctc cggaagaaat tctaagctcc aaagttaaaa tggcgagacg tggcgcatat  
55921 cagcagacgc cgtcactac ttgtggctcat tttaactctg tctgactccc ggcctctgtt  
55981 ttccaggaat cctgttttgc ggacacccag tggacagcga aaggagagag accaccacag  
56041 gcaaacggcc gaaagccgc cgactcgacg cccgcgcgtc tcaactccgc caataaattg  
56101 actacgcgtca ccaaccccat ctctgtctccc cgagccggcg acctgaggga gcgctgtccg  
56161 cttcccgcga gatgctcacg cgaaaaaacc actcttcagt gttctgagac actccgcagc  
56221 ttctggcctc ggtgagcggg tgaactaaac ccgggaaacc gctgaaagac cccagctctg  
56281 actagccctc tctgtagcgg cggccagcgg gagggcgtag cctgtagctt gggccacgtc  
56341 attctgttgc ccgggtcagc tcaacccccc gagggatag gaagagaaat gaagaagact  
56401 gctactccca gacctctcc tatcgaggac ccaatcaac aaccaccctg tgagcgcca  
56461 tcttctctcc cttctctccg cgcgagcgac agcaccgctg ggcgggtctt tggccgaaaa  
56521 gaggcgcga aaggaaggag tgcgcgatgc gcaggcgtc tggccgaaaa ggcctctgag  
56581 gatttgtagt tctttgaaca ggagtaacct ttgtattctg acSagaccaa gcggtatgat  
56641 aggaataatc tcttcagcta gcaacaaatg cataactcga tctcttatat gcgcatatca  
56701 agggcgcttc ctaggcaaaa tatcaaatg acagggagat gccttttagg cgtctttatt  
56761 acgttctctta ctacgtcccgc ttcccgcaga aggggtacaa aatacaaaa gcttcccggg  
56821 acagMtRett taagtaagag tgatagggcg tttaaaacga caggtataaa tgagactagc  
56881 taatgagata Ycttggaact ctctgaaaac attcacactc aaaaactttt tttaaacacg  
56941 atctctctac ccgcaaaaac actctcttgg ggagatattt cgttaactctt ggtaatttac  
57001 cgaggtctctg gttctctctg ccaagggtatg acaagctcga attctctatg gaggctgtc  
57061 agaaaaaac atctcctaac tcatctctgt tggcaaaaa actctctctg cgtttgtggg  
57121 actcagttcc ctgttttgtt ttgtgctgtt agtgggtggc caccctcagc ttgtgtggg  
57181 actctcaggt ccttcccaca tgagccctat ttattgttag ttcaacaact ggtattttg  
57241 ttctcaggc acatctttct gactcctctt ctgcccactg tctctctatg ttaagatcaa  
57301 ttaagggct catgtgatta gatcaggagc acctggataa tctccctatt ttaagatcaa  
57361 cttctggcgc cgcgcagtag ctacacttg ctacacttg taactctagc actttgggag cgtggggg  
57421 cgggattgcc tgagctcagg agttcgacaa agcctgggca acatgggtgac acctgtctc  
57481 tactaaaaa caaaaaatta gctggcggtt gtggcgctc cctgtatgc taactagctg  
57541 gagggctgag gcaggagaat cgtgtgaatc aggaagcga ggttgcatgt agccaagatc  
57601 acgccactgc actccagcct ggccacagaa tgagactgc ttccaaaaaa taataataa  
57661 taataataa aataataaaa taataaaaaa gaatagctgt ggtgtgtgat gtgctctgt  
57721 ttctcagact acttgggag ctgaggtaca gaactcgtt gaaccacaga gaacctgtg  
57781 cgagtgagcc gagattgcac cattgcactc cagcgtgggt gacaaagtta tctctatgt  
57841 caaaaaaaa aaaaaatca actctgcacg accgcagaa agaatcgaa ttctcatgt  
57901 catgcggcca ggggttatgt aggtcaggtc ctgtgggggt ttgggtgaga ggaattctg  
57961 ggggatattc tagaattctg cctgcacagc catgtaacc aaaaaaaa caaagaaaaa  
58021 tcttttggtt aatacaggat tgttaatac at ttgtgtat ataacctgca tggcctggga  
58081 ccccgaagt tagagtttaa gatgtagctt tggctgggtg ttgtgtgtt tgagaccagc  
58141 cccagcactt tgggagcgt aggcgggagc atcacctgag tgcggagtt ggtgtgtgtg  
58201 ctgaccacaa tggagaaact gcgtctctac tgaataacaa aatttagccg ggcgtgtgtg

58261 tgcctgectg taatcccagc tactcgggag gccgaggcag gagaatcgct tgaacctcggg  
58321 agccggagctg tctgggtgag cgagatcgct cgtcgttgca ctctagctcgt agcaacaaaga  
58381 tcttaaacctg ctcagaaaaa aaaaaaaagat gtagcttcag caccataaata  
58441 ttcttatagc gcaagtcatt taatttattt ggcccccagt ttatttgaaac ttaaaatgct  
58501 gatcatgtata tctgctttac ctatatttat ataaatgtaa aataatata tacaatacag  
58561 cagagataga gccctaaatg aggtaaatgt ggtggatgt tttttaacta actcttaagt  
58621 actatgtatata tatgttagtt attacagtta atatttttgt tgacagtgat ctacttcaac  
58681 cctctgcca ggaatgtatt atttctaaaa ggagaaaaac agagaagcag attatagaac  
58701 gctctggaag ccaagatcaa ggagcttaga tttagttcag cagggtattt gcagctgaag  
58841 gaggacttaa gttagagggcc tgggtcaaa gaaacaggcaa gaaagacatt tcttaaggag  
58861 gaatgggttt ctatggcaat gagagacaaa ggaggctact gagggacaagg ttaaaataat  
58921 taacatggga aatcagtaga acatagacca agacctggga tgcaggattg aaagaacagg  
58981 gtggagaaga gcttttgcaa tccactctca ttctccacag aggcacatca tgcgaaggat  
59041 ttccaccagt catagtccct cgctgttttg ggggtgtgta aagaatagaa actgaactta  
59101 atagacaatc tatcaaaaaa attaaagaa caaatctact gcaaacgtgW tacaattata  
59161 taccaaaagg cctagtgcct ttcttgtttt tgttggggag ggggttttgt gttaagctat  
59221 gctgatactt cagtgcctat ccactccctca ttatcatag taattatgag ctgcctgctc  
59281 atgtatagag ccagtgcat gctggaatca gtaatgcgc aggttgcaag aagcaatcaa  
59341 ttcaactcag tcttagaatt ctggatagag agtagaattg ttctcagaga acctaggggc  
59401 taattttgatt ctaagtagg tgtggctatg tagccacaag taatatagct ttgtagggtg  
59461 tagggaggac aagttaagca catattcttg agtcaagctt cctgcctctc tggccoaagt  
59521 cctcaattta tgcagctgag attctggata agtcaactct cctctgtctca gttttctgac  
59581 tggccaataa aaagggtgag gggcgtggaa tcaagtaaac ttatctccta taaagctatt  
59641 gggagagctt aagaggttaa tcatgtgaa gctctagaaa aatgccacag taatagtaga  
59701 gcctcaaaaaa ttaattatgc ttctcgcttt taatgaattt gcccttaatt attaaggttc  
59761 aggttttgca tctgcaaaat gaggcagttg aactcaattt tatctttggg tatctctccc  
59821 actaacattt tctactatag gatgttttt taaatttact tcaattgcaa tattttatgc  
59881 tctgatgcca ttatgcttta gttattatgg ctatattatt ataactctc tttagataaa  
59941 agtagagctg ctccaccwca ccgtctttac tttttgatg ttgttttact gccctctcaa  
60001 tgttttaaaat ttttttttta taaaaatgt gttttgcat ttgacatagt taaaaaggct  
60061 agagccaagt attacttagg aactgcgaaa gaaataacgt atcttttaga tggagagatt  
60121 tggctgtcat cacttttgct tcaactaacg tggacaacat tgcactatgt ttgctttaat  
60181 aggaagttcc ccagttttct ttctttcttt ctctttcttt Yttttttttt  
60241 tagacagagt ctctctccgt tggccagggt ggaatgcagt ggtgcatttt cggcttatgt  
60301 caacctctgc ctccctagt caagcagttc tctgcctca gccctccag ttagtgagag  
60361 tacaggctag agcaagcaag cctggctaat ttgtcattt tttagtagag tgggttttca  
60421 ccatgttgct caggctagtc ttgaactcct ggctcagaat gatccacca ccttgccatc  
60481 caaagtgct gggattacg gtatgagcca ctgcacctgg ccgRaagtt accaatttc  
60541 tgatcaaaac agacctaat tccagttcac aagaaattca ggggatagag taacaagttt  
60601 aaaaagctga ggaacaatc agtaaaatcc agaatgtgag acattctaca agacaactga  
60661 ctgagactct acaaaatgct aatgtcatgg aatggatgtg aatagaaatg atgaaaacag  
60721 attgatgaga catgacacat aaaattagta aaactgtatg ggaattctggt tcaagttggg  
60781 aaaaactata aaaaatttct tgggacattt ggagaaagt gcatattgag tgataataga  
60841 caacattgga gaatttttgc taaatttctt aggaattata atgacatgtt ggttaagtag  
60901 aagaacgtct ttattcttgg gagatgcat ccagagatt ttgggaatgaa atgttcaacc  
60961 ttactttctt cttttttgtt tgttttttgt tttagacac aattttgttc ttgtcaacca  
61021 ggtcggatga cagtgcagca gtcttggttt actgcaactc ctgcctcccg ggttcaagtg  
61081 attctctctgc ctgactctcc caagttagct ggaattatgg cgctgcacac gatgccacg  
61141 taatttttgc gtttttagta cagacgaggt ttaccatgt tggccagcgt tggctcgaac  
61201 tctgaccctc aggtgctcca tccacctggc ctcccaagt cgtgggatg taagctgag  
61261 ccaactatgcc cagcctcaat ttactttcaa atagtttatc tcaacaacaa aaaaatgtg  
61321 tgtgtgtgtg ttatagtga gagagtgta gagagatga aagggaagc aaaaatgtga  
61381 caattatttta agcttagatg ggggagtggt gactgatag gtgtttatgt aactatttgt  
61441 tcaacttttta tatatatag aaaaattatc ctttttttt tttagacagc aggtctcaact  
61501 gttgtgtcaa gctggagttg agtggatga tccagcgctca ctacagccc atctccacg  
61561 gctcgaagca tactcctacc tcaagctccc aagttagctc gactacaggc atgtgccacc  
61621 aggcctgctt aatttat ttt ttgtagagac aggtgtctca tatgttgccc aggcctgctc  
61681 tggactccc gactgaagt gctcctcgc ctgcatccc aagtgatgt agaccaactg  
61741 accagacctt gaaaattttt ataaataaaa tttaagaaga ggaaggtgtt agaatcaca  
61801 ttctccaatc tcagaaacct actcaaaaag tgcagtaatc agacacagtt ggtactgcca  
61861 taaggataga Ytcatagatc aatggaataa aatggagagt cctgaataat acccatatgt  
61921 ttctgggtga ctgatttttg acaagggtac caagacatc taataactat attgacatg  
61981 ttttttcaac aaatggtgct gggaaaactg atgtctaaat gcaaaagatc gcagttgga  
62041 ccttgctcca cgcctatac aaaaaataac tcaaaaatac tcaagacctt aatgtgaaa

62101	actaaaaacta	taaaactcttt	agaagaaaaac	acagggggtca	atcttctatga	tctttggattt
62161	ggcaaaagaat	tcttagatata	gatacaaaaaa	gcataaaaaaa	caaaaggaaaa	aataagaaaaa
62221	ttggaaYataa	tcaaaaagtta	aggttttittg	cttcaaaagga	catcatcaag	aaagtgaaggg
62281	gcggggcagt	gtggctcacg	ctgttaatcc	cagcaacttgg	ggagggcgag	gcaggcgagat
62341	cacagagatca	ggagatcagag	acgatctctgg	ctaactctgtt	gaaaccccat	ctctactaaaa
62401	aatacaaaaa	ttagccgggt	gtagtggcg	gcactgtlag	gcaagctctac	tacagaaggtc
62461	gagggcagag	agtgccgtga	accaggagag	cagaacttgt	agtgagccaa	gatttgcgca
62521	ctcgactcca	gcttggggcaa	cagagcgaga	ctcgtcttaa	aaaaaaaaaa	aaaaaaaaaa
62581	aaaagaaagt	aaaagtccca	agaccttccag	gcaaaaaaag	aaacttccat	taggctatgac
62641	atttcaggag	tttagagatt	acttccaaga	agctaaaggc	acaacgcaga	tctctctttt
62701	gacaaggtta	aatttggttac	tacagaagcc	ttgaaaacat	tatgtctaat	gaaagaagcc
62761	tgctcaaaaa	gaccagatat	tgtattatta	cctttatata	aaagtgttca	aataaaaaaa
62821	tctgtggggg	gtgttaggaa	aggtcaataa	ttgttaatgg	gtatgagttt	cttttggggg
62881	aggggttga	aattagtggt	gacgggtgtt	caactttgaa	tatactaaac	gctcttaatt
62941	tgcatacttt	aataaaggtga	ttttatgggt	cgtagaactac	atttcaataa	agctgtagaa
63001	gcaaaagtta	gtaaaaatga	aaagtaataa	aagtgaataa	tgataaaaaa	gtataataata
63061	gtatctcttc	atttttttcc	aaatgaattt	tagaattatt	ttgtgcatct	tattgaaaaa
63121	tatctgtagg	caatttggatt	ggcgttgaat	tgctcataaa	ttaatcttga	gagagccaga
63181	attctttacaa	ttttatgtct	tctcactcaa	gaacactgg	tttggttttK	gttttttagt
63241	tgaggtcttt	ctctgttgcc	tgggctggag	tgcaaggagc	gtcttggctt	ccctacaacc
63301	tctgcctctc	gggttcaaga	gattctcctg	cctcagcctc	ctgagtgtct	gagatatacg
63361	gcagccacca	tccacctctg	ctaatttttg	tatttttagt	agagacgggg	tttccgcat
63421	ttggccagcg	tggtctcgaa	ccctggaact	caagtgtact	gctctgctKt	ggctcccaaa
63481	ctgtgtggat	tacaggcggt	agccactgct	tttggccagg	tgataaaaaa	ttttatttaac
63541	ttattaaaaa	cttctctcat	atctttcaat	aaacttttga	agttgtcttt	atatagattt
63601	tacacactct	tcaggatatt	aatttttggt	tttcaatttt	aaaaatcatga	atgcaactct
63661	tgtttttata	tgattactaa	atagaacagt	aatttttatt	aaagtaagta	aaaaattgtc
63721	gcaaaccttc	agtcatttta	taggcagaat	caattttttt	aatttttatt	ttttattttt
63781	tacaggtatg	atcaaccagg	tacagaatca	gttttttggc	tgtagatgat	taagagatta
63841	gggtcaactt	gtctttgggt	agatgtgctc	ctctatcttc	ttctgggaat	aaacagttct
63901	ttgggttgat	gagtaggaaa	gcagtttttg	tgtagcgtac	ttccctccat	gagttctcag
63961	actctctgtt	taaaagcact	acgaaacctg	ctgaccccaa	ttctttcccg	attttaatatt
64021	tctttttgaa	agttggctct	aaccttaagg	gttgtgaaaa	cacaaggatc	cagttgtggt
64081	gaagctttag	aactaagaaa	accattctct	ttctatcaaa	acagctctgg	agggggaacg
64141	tttgcctctt	gcccagagcc	ctgtcatatt	ttttgatcat	tttcatctct	atcnaatctg
64201	ctgttggccc	agaaggcagc	catctgatag	ctgaagtga	tgctcatatt	tcttctctct
64261	cttctccagt	aacaagatat	tccaccagaa	tatgctgtt	tgtaggctaa	ggctgtgatt
64321	taataaccct	aaggagataa	acactgggat	aataataatg	gttattattt	attgatagc
64381	ttctctatgc	caatttatcc	cacttatctc	ttgtaacagc	cctgtgagta	gttaactctt
64441	ttttatccca	ctttacagat	aagaaaatgg	agggctagat	atgtaagctg	cccaagatga
64501	caaatagaaa	gtgactaatc	tggaacttaa	accacagaaa	catgatctat	cttttttttt
64561	ttttgagaaa	gggtctcaat	ctatgcgcca	ggctgggggt	caggggttga	ataccggctt
64621	atttgagcct	tgacttctgt	ggctcagctc	atctctctgc	ctcagctctc	cagatagctg
64681	ggactacagg	cgaggccccc	cagcccccag	taatttttct	atttttttga	gagatggggg
64741	cccactattt	tgcccaggct	ggcttgaac	tgctgggctc	aaagcaatct	cctgcctcag
64801	cttcccaagg	tgctgggatt	acagggtgtga	gccactgtgc	cggcccgaga	gagatcagca
64861	ttaaccaata	tgcttctcaa	gtgaagaatt	ctacatttct	agaagcatct	gagttcttct
64921	gacctgagtt	caagtctgtc	cttctctccc	actttactga	accttcagtg	gagttcttct
64981	cctcagaagg	ctgtgaggt	gccatggggt	aatgaatgta	aaagtgttaa	gtgcttggct
65041	gacactgaggt	actaaaataa	tatcaattta	ttcttcaaaa	gtgtataaaa	atgtatcaga
65101	tcttcttctc	gtactaaaac	ctgtactaat	ttgggggaaat	tcaaaagttaa	agtcctgata
65161	gtcttgcctc	aaaaactaat	gggctgggag	gggcaacagt	cttcaacctc	gtatctccag
65221	cactttggga	ggctgaggtg	ggcagatcat	gagatcagga	gtctgagacc	agcctggcta
65281	acatattgaa	ccccatctct	actaaaaata	caaaaaatga	gtctgggcat	gtctgggggt
65341	cctgtaatcc	cagctactcg	ggaggctgag	gcaggagaat	tgcttgaacc	cgaggagcag
65401	aggttgaggt	gagctgagat	cagcccatgt	caactccagc	tggggcaaaa	gagcgaactt
65461	ctccatctca	aaaaacaaaa	acaaaaaatc	aaaaacaaat	ggctgtgata	cttgggactt
65521	gtgggttgaa	tctggccccc	atacatattc	tgttttgtct	tcaactgtct	taaaactttt
65581	gaattagttc	cagacattta	aaagtgtgga	gattttggcc	gggcatgtgt	gctcacactt
65641	ctaatctcag	caacttggga	ggccagagcc	agtggtccac	ttgaggtcag	tgacttcaga
65701	ccagcgttgg	caacatgggt	aagccccatc	tMtactaaaa	atacaaaaat	tacggcagaa
65761	tggttggcgc	acctgtaatc	ccagctacat	gggaggtcga	gcaggagaaa	tcgcttgaa
65821	ccaggaggtg	gaggttgcag	cgagctgaga	tcgtaccact	gtactccagc	ctgggcgaca
65881	gagtgagact	ctatctcaaa	aaaaataaaa	ataatataaa	ataagtttgg	agattttgat

65941 ctgtttccacct gcttttgaac ttctaccaca gacaccttct caagggggRt ggtgtctctg  
66001 gcgatctgcac cactggcaga ggaaggatata ctgaaccacag lcaattggcca agctccctga  
66061 tattagttcgt actacatatct ctgacctctt ctgggtccac cagacatacaca tglaccttgg  
66121 ggaccaccctt gcaaaatgcc atcaacacag gctctcaaa tagggacattt tacaccaatc  
66181 accagtttta gtaacacaa taacataggg ctattccctc gaatggaagag ctgagaatgt  
66241 gttagtttta aatatagaat cagagtcagt gtggcaggtg gaatttaccg acttaaaatc  
66301 caaaaaacca cacttaaatg cctgtttctc acaaaagaca gcttaaaatc cctccacagg  
66361 caataaagag aagcctccat aggttgagat gtttctctgg agggaaatct gaaggatga  
66421 gggagatga ctctagctcc ttccacacagg aggtatagaa ttcaactctaa gggcaggagc  
66481 cagatcgagt taataaggaaa aataaattcc ttcatagtg gtatgtctaa gaacccaca  
66541 cccaattttt tttagggata tctcacaag gatagagcag gtgtgggtct gacaggtggg  
66601 actgacttaa ggcataccct gctcacaac agaagcccat tccaccagc gatctgagag  
66661 atttgccaga gctgggtga aggttgagga tgaaggagc tttaacttgt gattctgaag  
66721 ttaaatccaa tctacagta tgaactctgc tccaaaccac aataccgct cttaaccact  
66781 aggatagcta taaccaaaat gacataataa gagttagtga ggaatcgag aaactggaa  
66841 gcttgtacat tgacaggtgg aatgtaaaat ggtgcagcca agtgagataa ctactggga  
66901 tcttctcaaa actttaaacg gagagttacc atatgatag gtttggtctg gtgtccccc  
66961 ccaaatctca tttgtaattg taattcccaa tgttggggga ggaacctgat gggaggtgat  
67021 tggatgacg gggcagattt ccccttgcct gttctatgat agtaggtctt gctgtctat  
67081 tagatgtagt agatctgat ggtttaaaag tgtgtgcac ttcccctctt gatctttctt  
67141 ctgccaccat gtgaagaagg tctgttgcct cccctttggc ttctgccatg attgtaaaat  
67201 tcttgaaacc tccagtcac gctctgtta aaactcgaga tctgtgagtc acttaaaact  
67261 cataccttag tacaccatat gaccacgcaa ttccattct tttttttttt tttttttttt  
67321 tttttttttg gacaggtct aactctgtca cccaggtctg agtgcagtg ttcaatcaca  
67381 gctcaattgt ccttcaacct ctgtgtctga agtcactct ccaactcagc ctcttgagta  
67441 ctgtgggact cagggccatg ccaacctgcc ttggtatttt ttttttttaa tttttgttaa  
67501 agacagagtc tccctatgt ttccaggtct gttctcaaat ccttgagacta agcaattgt  
67561 cagctctcat tcccaaaagt actgaggggg attataggt tgagccctc cctccagcct  
67621 ctactcttaa tatgatcaag agaaatgaaa acacatgttc acacaaaaac ttatagcga  
67681 attttcatag cagcattata atagctgaaa tgtggaagca gctcaaatg ggaagtgga  
67741 gatgaatgga taacaaaaag gtggtatatc cacacattga aataattact agcaataaga  
67801 agaaatgaag tactgatagc ggatacaact tacacgaact ttgaaaaact agactgaag  
67861 gaaagaaacc agtcacaaag accagatatt atataattcc attatttga attgttcaac  
67921 atagggcata ccatcggaat gcagattagc agttgctctg ggaatgggct ttctttttg  
67981 ggcaggaagc agggactgag aaaggaaca aaacagcagt ggaatgggag ttctttttg  
68041 agtatgtgaa gttaaaaagt tctaaaatta tataatggt atggtttgac aaactgtga  
68101 ataactatga aaaccactga attgtacaca ttgaaagggt gaatttttac gactgtgaat  
68161 tatactctaa lacatctgtg ttttggccag cgtgtgtggc gtgttggctc aaactgttaa  
68221 tccagcactt ttggagggcc gaggcagcgc gatcatttga ggtcaggagt ttgagaccag  
68281 cctgatcaac atgggaaaaa cccgtctcta ctaaaaatac aaaaattagc aggcattggt  
68341 gcacacacact gtaatccag ctactcggga ggctgagga agggaaatgc ttgaaaccag  
68401 ggcgcagagg ttgtggtgag ctgagatcgt gccattgcat tccagccttg gcacaaagag  
68461 caaaactctg tctcaaaaaa aaaaaaaat ctgtttttta aaaaataatc ttgtctcact  
68521 ttctcctaga taccctaa gataaccagga accacaatgg acttttccct ttctctccg  
68581 tatttttaaa aaataattcag ccaggcatgg ttgtgcacct gtaatccag ccacgtggga  
68641 ggcgtgagtg ggaatattgc ttgaacctgg gaggtggagg ctgcagatgc tcagatcgc  
68701 accactgcac tccagcctaa gcaacagagc gagactccat ctacaaaca taatttata  
68761 atgagctcct ctttaacttaa ttcttttaa atgttttaata gctgttataa taatttata  
68821 aattattatt aKtagtattt gagacagagt ctcaactctat tgcagggtct gatcacatga  
68881 gtgtgatttt ggctcactgc aacctctgcc tctgtggctc aagtgtattc cagccttag  
68941 cctectgaat agctgggagt acaggcatgc accacaatgc cgggtctaat ttgtgtttt  
68961 tagcagagat ggggtttcat catgttgccc aggetggtct cgaactcctg tctctaatg  
69061 atctgctcgc ctgagtcctc aaaaagtcgt ggattacag attaggcac agtgcctggc  
69121 caaaaaaatt ttttttctt ttttttttt ttcaagagaca ggtctaacat ggtctaccca  
69181 ggctaaagtc cagtgacacc atcatgactc tagataactg actgtctgct cgatttctg ggtcacaatg  
69241 atctctccac ctgagcctcc tagataactg ggaaccagag cagctaacac atgcacagag  
69301 taggtgtatt agtctgttct catgctgcta ataaagacat atccgagact tgcagatga  
69361 taagggaag aggtttaatg gactcacagt tccaatggc ttgtgagacc ttgcaggcaa  
69421 ggtgggaagt gaatgagga caaagtcaag tcttacaatg ttgcaggcaa gatgtgtcc  
69481 aggggaattc cgttttataa aacctcaga gctgttgaaa ctattttact atcatgagag  
69541 cagcagagga aagaactgcc ccatgattc aatcactccc cggggggtcc ccccatgag  
69601 atgtgggaat tatgggagct acaattgaca atgagatttg ggtgaggaac cagccaaacc  
69661 atactcagat ggtttttttt ctttgataga gatgaggtct cactatgttg cccaggtctg  
69721 tcttgaactc ctgggctcaa gcaattctcc cactttggat tcccaaaagt cgtggattac



69781 aagtggtgaac cattgcgcct ggcctcatgg cctagttaac tgcctctttt ttttcccgag  
69841 atagatgcttt ttctgtgcac ccaggctgga gtgcagtggt tccactgcag tccactgcag  
69901 cctgycctctc ttgggttcaag cgattctctct gcctcagcct ccagagtaag ctgggatatac  
69961 aggcacccgc cactgcacct ggcctatttt tttattttta gttagatagg gttttaccac  
70021 ggttggggcgg gctgtgtcttg aactcctgac ctacagtgat ccccccctgcc tctgctctccc  
70081 aaagtgcctgg ttattcaaggc gtgagccacc atgcctggct agttaactctc tggacctcaa  
70141 ttctctcagct tgaataatga ggataataac agtcccttcc tcatagggta tatgttaaatc  
70201 taaaatgacc acagatatga actatgttta gaacagtacc tagcacataa agtaccctgt  
70261 attagtatttt gtaataatta tttatatatt aatatattgt atacatatata atagtcatatt  
70321 atataccccc taggatagta tttagcacat agtaagtgcc ctatRtaagt ctatgtatttt  
70381 aatatatttca tttatttact atttttatta gaggttgataa tggtttaacct gtcccaaggct  
70441 ttatctgttc ttgctgttgat gaacatttgag ctgtttccag tttagaggta caaatgcgta  
70501 atgctgctctt gaataattgt gcattcatttt ctctagggta taccattagg agtggatgac  
70561 tgaggcgaag ggtatgtttta cctccatatt tacagatatg ccaaacctgtt tcccaaatgc  
70621 gatgatctcaa tttacattcc caccctcaggt aagaacctct ccaatattctg ataggctcat  
70681 cttttaaaat ttaccctggg tatgtaatga cagatcaaat tttatttttg attttctggg  
70741 ttactgtatga ggaggagcac ctttgatttc ctctctgtga cattctctgtt caagcctctt  
70801 tgcattttac tactgagtta tactgtctttt tctgtttgct tttgtagaat tcttcaactg  
70861 ttctgggtgac tgagtaatttc tttatNtatt tcaataacct cccatttctt tctgtacctt  
70921 tactctgcga ttgggtatctt ttgtttttgt ttgtttttgt agatggaggtt gttttttgt  
70981 tgccccagcgt ggagtgcatt ggtgctatct cagttcactg caacctctgc ctctccatgt  
71041 caagcgattc tctgcctca gccctcctca tagctgggat tctagggtgc tggccacagg  
71101 ccgggcaaat tttttgtatt tttagttagg acgggggttc accatgttgg ccagggtggt  
71161 ctgcgaacct agaccttaaa caatccaccc gccctggcct cccaaattgc ttccaaatgc  
71221 ggtgtgagcc accatgcctg gccccagtggt tatcttttga tgaacaaaaa aatttaattc  
71281 agaaagtaagc aaatttatga atctttttca ttgtagttag ttgtcttttt ttctgttcca  
71341 agaaattctct cctctgctctg aggtcttattc ttgtaaaaag tttatagtatt ttctttttac  
71401 actttagtgca ttccacctg gaattgaata ttgtgcacga tctgatattg acttgaggcg  
71461 aatttcaattt tttctacatc agatgtcctaa ttttctcagc accattttatt aaaaagcctg  
71521 ttctttctctc actcatctgc aggcacatctc ttgtcatgtag caagtgttat ttgtcaaatg  
71581 tcatgtatttg atagttttga gggctctttaa ttgtttctct tttacaggaa ctctctccac  
71641 ctgtgtctctc ttattcaata gtgtttttgt ttgtactctt ttgaaatagg ttaattagg  
71701 tccaccataaa acaacaaga aacaaacaa aaccctaaaa lggagcaaaa ccattttccc  
71761 agtgcataaa actctcaaga tctaactgaa ttgtagtgc cggactcttt tcttttttca  
71821 cctggacatg aacgaaagc atgtagattt aggttctctgc aacagctccc tcatcttagg  
71881 cacattagta taggggtgaa aacaacactc acagccgggt acagccggct atgcctgtaa  
71941 tccagcacct ttggagggtc gaagcagggt gatcacctga ggtcaggagt tccagaccag  
72001 cctgcacaac atggtgaaac cccgtctcta ctaaaaatac aaaaattagg ttggcatagt  
72061 ggctaacacc ttgtaatccca gctacttggg aggcctgagga aggagaattg ctgtaacctg  
72121 ggaggcgagg gttgcagtga accgagatcg ggccattgca ctccaggctg ggtgtcaaga  
72181 gtgaaactcc atcatcaccc acatacacac acacacacac acacacacac acacacacac  
72241 acacacacac aggcctggag cagtgtgctca cgcctgtaat cccagcactg ttgaggcgcc  
72301 aggtggggcag atcacagggt caggagtttg agaccagcct ggccaaatac tgaacacccc  
72361 acttctacta aaaaatacaa aaattagctg caagtgttgg ttggcgccgtg taattccagc  
72421 tactcgggag actgaggcag gagaattagt tgaacccagg aggcagaggt gtcagtgagc  
72481 cgagattgtg ccaattgcact ccaagcctggg cgacaagagt gaaattccat ctcaaaaata  
72541 aaaaagaaaa caacactcag aaggcgagag ccaagagaat cactaagaac agagtgcagc  
72601 aatgtctggg tctctcttac ctctcaagt accttattca actctgttaa atcggtttaa  
72661 ctgttatttg caaacactg aaagcatcct aacttattca taagctattct taaataattg  
72721 ttctgctgca gaagcaagta aaagacctc ttctgtgaaa ggaataatac ttgtcacaaa  
72781 ctatttaattt tattctatgg gggccaaaca tgataataa atataataa ggtgatgttg  
72841 atagtataag ataaacaatta caaagccagt cctatgtcta aggtattttac agataataa  
72901 tcatataact agtattctta ttcaacggcc atgatctccc agtgttactc tttttagaga  
72961 ggaggaaaaca ggtccagaga tcttaagtaa ctctctcagt atcacacagt ttttagaata  
73021 caaaggtagc cttgactggc tccagcacac ttatacaatt cttcttattc taattattat  
73081 tattattttt ttgtttttga gaYggagct ctctctcagt atcggttttg ctgagcagtt  
73141 cagcatcttg gctcactgca agctcgcctc cctgggttaa ccgcatctct ctgctcagc  
73201 cctctgagta ctggggacta caggcatctg cccacttccc cccgtaattt ttgtatttt  
73261 tagtaaggac agggtttcac cctgttagcc agggagtgtt cctcatctcc tgcactctg  
73321 actcacctcc cctggcctcc caaagtctgct ggatagcacc ctgagccacc ggcgcccgc  
73381 cttaacttat tattaatagt agcaaatgta aaactccag gaggccaggt ggtgtgactt  
73441 atgcctataa tctcaacacc ttgggagggc aaggcaggtg gatccactga ggttggaagt  
73501 ttgagaccag cctggcgac atggtgaaac cctctgtgta ctaaaaataa aaaaatagcc  
73561 aggtgtgttg gcaactgcct gtaatccag cctatggga gctgagaca gctgaatgac

73621	ttgaaccacg	gaggcaggtg	ttgcaatgag	ctgagatagt	gctactgcac	tcacgcccgg
73681	gagctgtttc	aaaaaaacaa	taacaaagac	aaacaaacaa	acttccagg	caggcgttgt
73741	gggtctagct	tttaactccta	gcactttggg	aggctgaggg	aggaggtctg	tttagcccca
73801	ggagttccaa	actagccctg	tcacaaatag	ggagccccat	ctctataaac	aaactttctc
73861	ccaagtctgt	ctcagtgatg	tgagtgagtt	ctcacaagat	ctgatgtgtt	tataaggggt
73921	ctctccccc	tcacttgatg	cttctctctc	ctgccacott	atgaagaagg	tgctctgggt
73981	ccccctttgc	ttctgcaatg	atataagtt	ttctgagggc	accocagcca	tgctgaaactg
74041	tgagtcacgt	aaactctttt	cccttataaa	ttaccacgtg	atttataaac	caactcaagg
74101	tgagtgttga	tagcagcatg	aaatggaggt	aaalacagct	ccctaaagtgc	tggtgattaca
74161	ggcatagacc	accacacgca	gcctgcattt	tttcatagat	aaagtctttc	gggtattttc
74221	attgacacat	tggaattttc	agaaagatg	agaaagaaaa	ttttaaatalc	ccatlaattct
74281	actctgcaga	gataatcaat	attatcattt	tatagtatat	cccttgagct	ttttaaaataa
74341	tatatgcaat	agaaatccag	ctgtgaaat	tgttttgtaa	tcctgtctctc	tttttttttt
74401	tgatatggag	tcatactgtc	actcaagctg	gagtgcaagt	tcctcaactc	agcttactgc
74461	agcctccacc	tcocaggttc	aagcgaattc	ccctgcctcag	ccctcccaagt	agctgggatt
74521	ataggccgct	accaccacac	ctggcttaatt	tttgtattat	tggtgcgac	tttctttcgc
74581	catgttgccg	aggctgtgct	cgaactccgt	accctcaagt	atctgccacc	tttggtccctc
74641	caaaagtgcg	ggattacagg	tgtaggcac	tgacactggc	ctgctttctt	aaagatgat
74701	ttttttctagt	cattaaaaat	tcataaatat	tattattaat	gtttgcgaatg	gtattttagt
74761	ttctaaactt	tacttaatac	gtgctctatt	ttcactaaaa	tttctgttgc	atacacattc
74821	tgaaaaactt	ttttttcaga	caaacactgat	ttttttaaaa	gtagaattac	tttataaagg
74881	aaataataca	tttacaagtg	tttagtacta	cacaactgct	ctccagaagg	ctcatatcca
74941	cttctgttcc	taccattagt	gaggtagaat	tcctcttttc	ccaaaactct	actctctctg
75001	aaagtttttt	tttttttttt	taataagaca	tggtcttgct	ctgtacccca	tgctggagtg
75061	cagtgtgtct	atctcagcta	actgcagctt	tgacctcttg	ggctcaagcg	atctcccccac
75121	attctactct	agcctcctag	gtagctggaa	ctacaggtgc	actccacact	cgctgtcgct
75181	ttttttcttt	gtttgtttgt	tgagcccatc	gattcacaat	cttccacctg	gtctctccct
75241	tgacacgtgc	agattatggg	gattatgggg	attacaaatc	aaggtggaat	tttggtgggg
75301	acacaaagcc	laaccataatc	agagggcagg	gtggggagat	tgcttgagcc	taggagtaca
75361	agaccagcct	gagcaacaga	cgaagattgt	ccacaaaaaa	aggaanaaat	atttataaaa
75421	gacaaagtaa	acaaaaaata	aaaaaagtta	ttagatgaat	gtgtggccct	atgctgtctg
75481	ttccagctac	ttgggaggtc	gaggtggggg	gatcaactga	gctgtggagg	tcaaggtctg
75541	agtgcagctg	gattgcaccca	ctgcactcca	gcctgagcaa	gagagcaga	ccctctcaaa
75601	aaaaaaaata	aaaaaaagaa	agaaaaatca	aaacttttca	agagctagaa	aaagagaana
75661	atgagatgaa	ctgccccagc	ctttaattat	tgtaacacaa	aaagcacaat	tgctcttgat
75721	attcctctctc	ctcactccac	attattttta	agcaaatatc	aaacactata	tttcaattgt
75781	aaatattttg	gcattgtttc	ctaaccaata	agggtcctct	tttattttta	tatttttttg
75841	agacaggggt	tcactcttgc	ttaggggtga	gtatagtggt	gtgatcatag	ctcactgcac
75901	ccctgaactc	ctgggctcaa	gtaactctcc	tgctgtggcc	tcacaaagta	ctggggagct
75961	aggctYgtgc	caccacgtct	ggctaatttt	taattttttg	ttgagatagg	atcttgctat
76021	gtgtccacgc	ctggtctcaa	actcctgagc	tcaagcagtc	ctcctgcctt	gaactcccaa
76081	cgtgttagga	ttacaggcat	gagccactgc	acctggccaa	gggctctctt	ttttaaaaaa
76141	cacaaaccaa	gtaccatgat	cacatctaaa	tttaaaaatg	taattctttt	aaatatctta
76201	aaataccaaa	taataattcac	attgcctctg	ctgtgttttt	tatgttttta	tataatttat
76261	ctcttttact	taYaggtccc	ccctccatct	cttttttttc	atgtaaatat	ttgtttgaag
76321	aattgggtcta	ttgcctgtgt	agttttgcat	agccttggtt	ttgttatatt	tgatgcata
76381	tggtgtttaa	catgtccctc	atctgtctgt	ttcctgtaaa	ttcttcatat	gaactagagg
76441	cttgatcatg	ctcagagttt	tttgttttgt	ttgtgcagta	agggcgagga	agggcagga
76501	aaacttatca	aggggaggca	tttacttcta	ctgaagatg	ccagtgctct	atggtcttct
76561	ttgtgatgtt	agcaactggt	tatgatcaat	gccacaccca	ctcatgtcta	tggtgtgcaa
76621	attggcgatg	ttcttatctc	attgtttctt	ttcactttag	tagttagaat	gcctccattat
76681	ccaagagaat	tttttcttta	actattttag	tgctcttttg	aaagagcac	aaaggaaga
76741	gaaataaat	acctgattct	ttctctttta	tgaccagttt	taaaaaaat	taggtgcggc
76801	cgggtgcggt	gcctcacgct	tgtaatccca	gcaactttgg	aaagcgaggc	gggcgataca
76861	cgaggtcagg	agatcgagac	catcccgctc	aaaacggtag	aaacccgctg	ctactaaaaa
76921	tacaaaaaat	tagccggggc	tagtgccggg	cgcccttagt	ccagactact	tgtagggctg
76981	agggcaggga	atggcgtgaa	ccggggaggg	ggagcttgca	gtgagcgagg	atccccccac
77041	tgactctcag	ctggggcgac	agagcgagac	tcctgtccaa	aaaaaaaata	catccagtaa
77101	aaaaaaaata	aaaaaaaata	ggtgctttcc	tagcatgctt	cagtggtgac	tgtataaatt
77161	ttttttgtta	cgttataaac	tcacagagtt	aaataaactt	aatatgctt	catccagtaa
77221	tggtttttgc	ctattgatgc	tcaaaattatc	ctccttttag	caagtagaag	gttctccatg
77281	ttagctcttc	agtccttttt	aaaccaaacc	catatgctct	tgatggcttc	ctgctcttct
77341	gggtataaag	gtattccagc	ctggggcggg	tgactcatgc	ctgtaatccc	agcaatttgg
77401	gagggcagg	ctggtggatc	acctgaggtc	aggagtttga	gactcagctg	gccaaactgg

77461	tgaaccccg	tccttgctaa	aaatacaaaa	aaattagctg	ggcctgggtg	tggcgcgcta
77521	taatcccgag	tacttgggag	ctctggagcg	gagaattgct	tgaatccagc	aggtgacagt
77581	gagctgagat	tgacccattg	cactccagcc	tgcccaacaa	ggcacaacat	ccatctcaaa
77641	aaaaaaaaa	aaaaggtatt	ccaggatcat	gtgtacatt	tgcataccca	gacttgggat
77701	gactcatttc	tgtaagagcg	cctggtttct	tttagtgggg	caatttgcac	gtataaagtg
77761	gtgaagtga	gtatttttaa	aagtaagaaa	taacagatac	taattgaatg	ataacagtac
77821	aaaaatgata	gaactatgaa	ggtattgcag	cccaagaccc	catttagtag	ggggagtagc
77881	ctctcaacat	taacagagcc	tgaaggaaat	aatgcattct	aatacatatc	cacacatcac
77941	caYggatgctg	agccttagtc	cccaagaaaa	ccttaaacct	tatttagcaa	acctatttta
78001	gaagtgatg	taaaaacattt	tttaaagttt	catatatgga	accttagctt	ctctggaaaa
78061	tttagtgttaa	tagaatgctt	catttgctca	ggtttatggt	gcagaaaaag	atcatattagg
78121	gagcacgtta	aaatataggt	gtctggacca	ggcaatggtg	gctcacgctt	gtaatcccgag
78181	cactttggga	ggccaaggca	ggaggatcac	ctgaggtcag	gagttcgaga	ccagccttga
78241	caacacaggg	agaccccatc	tttcaaaaaa	ataaaaaaaa	ttagctaggt	ggcacactcc
78301	tgtggtccta	gctacttggt	aagctgaggt	aggaggatcg	cttgaggatg	ggggttgagg
78361	cttaggtgag	gtgtgatggt	accaccacac	tccagcctgg	tccagtagagc	aacctgtgct
78421	caaaaaagaa	aaaaaaaaaa	aagtatgga	ttctgggtcc	cactcYagac	aacttgaatc
78481	tgggtgtgag	gcttcgaatt	tgtattttag	taagccttcc	aaatgatttt	gaagtgagta
78541	gtctcaaaaa	ctaacttttg	aggaacattg	ccagactgac	tgatagaaag	agagagagtg
78601	ataagggaaa	gaggaggaaa	aattgctttaa	aagagtgtat	cttggttggg	ccagatctgc
78661	ccctgtctga	atcctggcac	tttgggagcg	tgaaggcggt	ggatcacttg	agggccagag
78721	tttgagacaa	gcttgcccaa	catggagaaa	ttcttactaa	ataacaaaa	aaaaaaagaa
78781	aaaaaaagaa	gtaaaaaaaa	gagtgtatct	tgacgcatat	ttctgaggat	tgaaacgatg
78841	tttaaaaaat	tagtgattat	Yggccgggcg	cRgtggctca	cgctgtgaa	ccagatcatt
78901	tggggagccg	agggggggcg	atcacgaggt	caggagatcg	agaccatctt	gggcaacagc
78961	gtgaaacccc	gtctctacta	aaaaatacaa	aaaaaaatta	ggcgggcgct	gaactccagc
79021	ccgttagtctc	cagctacttg	ggaggtctag	gcaggagaat	ggcgtgaacc	cgggagggcg
79081	agcttgcagt	gagccagagt	cgcgcactgc	cactccagcc	tggggagcac	agcgagactg
79141	cgctctcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaatta	gtgatatacg
79201	gccagggcca	gtggctcacg	cctgtaatcc	cagcattttc	ggaggtctag	cgaggtggat
79261	cacttgaggt	cagtgtttcg	agaccaggtc	ggccaacatg	ggcaaacctt	gtctctacta
79321	aaaaatacag	aattagctgg	gtgtgttggc	gtgtgctgtg	aattctagct	acttggagag
79381	ttgagcgagg	agaattgctt	gaaccagga	agcagaggtt	gcagttagcc	gagatctgcg
79441	cactgccctc	cagcctgggt	gacagagtga	tactctgtct	caaaaggaaa	aaaaaatctg
79501	tcatctctct	tttctgattt	aaaaggattt	tgttctctca	tggaaaaactt	taaaaaatagg
79561	agcaagcaaa	aaaaataaaa	aaaatttgta	atcttaccac	tccggagacaa	tgttctattt
79621	aagtgtggtg	atattcttcc	agacttctaa	aaggaattgc	aattgtctct	ttttagtagc
79681	acaccactca	ggaacttgct	ggggcctctc	atctagcaag	tgggagagtc	cacactggaa
79741	ccagctgctc	gaatcctgaa	cttcatgtgt	ctccattgta	tttttttatt	ttatttttaa
79801	aaatttaatta	tttatttttg	agacaaggct	ctgttKccca	ggcgtgtgtg	tgcagtggca
79861	tgatcacagc	tcatggtagc	ctccatctcc	caggtctcaag	cgatctctcc	accttggcct
79921	ccaaagtgc	tggtattaca	ggatgatgcc	acgcgcccc	atgtgctttt	atctgttaacg
79981	gaacagagag	gtccacgccc	actcctagca	ttactttaca	cacatgtctc	ctgtttaaag
80041	acaaactcac	acaaatttta	gtttaaagat	cttaattggt	ttttatttgc	aattatagaa
80101	tcagccaaca	cctcattcta	gaaaaataga	tgaatataga	gtgttctgat	gagctgcaga
80161	ggagaggttg	gctttatagg	tagaagaggg	gcgaagaaag	caaaaattggg	ggcgggcgga
80221	tggtctcagc	ctgtaatacc	agcaactctg	gagcttaggg	ggcagacagt	acctgagctc
80281	aggagtctga	gaccaacctg	accaacatgg	tgaaccctgt	ttcttaccac	taattcaaaa
80341	atagctgctg	gtgtatgggt	cacacctgta	ttcccaagta	ctcgggaggg	taggcgagca
80401	gaatcacctg	aacctgggag	acagaaagtt	cagtgagctg	agattgtgct	agtgcatctc
80461	aggctgggtg	atagagttag	acctgtctct	aaaaaaaaaa	aaaaaaaaaa	aaaaaaagag
80521	gaagacaaaa	acagagaaca	aaatggagat	tggttgtttt	caaatgtactt	ctcttaagga
80581	gttaaaacag	aggtgattac	tttattatgc	tgaactcggt	tgaactgaa	ctctgttttg
80641	gttttggctt	aaaaagagag	tggagttctg	ctttttggcc	aggtgtgctt	ggaaactctg
80701	acctcaagca	agtcocccac	cttggccttc	caaatgtgct	ggattacaga	agtgaaaccac
80761	catgcccagc	tggacccctc	tattttgttt	tgttttttta	ggaaaacagg	accttaagga
80821	tttcaacttg	tggcacttag	cacatgtgac	ttccattctg	tttggccttg	ctctatgggg
80881	ccagtgctag	gaggtgggtc	caaatgaYgg	ccctccatta	actttgtctc	acaccggcca
80941	tcacttgcta	aattaaactc	acaattacaa	caaaaaataa	tgtgcctact	actgaactat
81001	ggccttctca	gtagggcag	aacatcatct	tatttatatc	agtatccttg	gaacacagcc
81061	tgtgtcttaa	ctgtggctca	atacctggaa	cacagtaaat	acacaaaaaa	tgtactctct
81121	ggcgaggggg	caatggagtg	aaaacaacta	ttctcattg	ccctccacag	actcgaattt
81181	cagccatctc	agacttcaaa	actgacttag	ttttctcccc	aaaaaacagt	tgcagactct
81241	ccaatgtcca	agaaatagaa	tatccttaaac	gaatttagggc	agaaacaaaa	atcccaaatc

81301	cacatattctt	catttgaag	tgggaagctaa	acatcagata	tctatggaca	ttaggatgac
81361	aactgtagat	gcgggagact	actagagctcg	gggaggagag	agggggcaag	ttagtgaaaa
81421	ctgttgggta	cttgggtgac	gggatcattc	ataccgccata	ccctcagcatc	acgcagatata
81481	cccaggtaac	aaacctgcac	atgtacc5c	taaaactaaa	atccagatttg	aaaaaagaag
81541	aaaaaaagaa	agaaacagaa	tatcttctctg	aaagccttat	aggggaaact	gcagagagaaa
81601	gcaaatcttc	ttagcatctt	accaccttac	ctaagcaaca	ggccccttga	caataaagtt
81661	atgaccggcg	tcagggtgtg	gattgaaaag	ccagcgcatc	gtgctgcaat	aaagcgtcac
81721	caggtttaga	acaaaaagcc	tggctgtctc	cttttctctc	ttagcctgcat	aaaacagctca
81781	tacttggatt	ctgagtgttt	gaaacaaaag	catcacaggag	gaataaagtt	ggattttaact
81841	gacactcttt	tggaaatgaa	ataccaatca	aaatgaagaa	ataatgtttt	gacttctcaat
81901	acaaatctact	ttaggtttatg	agagtgaaaa	gtctgatgaa	atacatagtt	ataaaaatgt
81961	gaaaaacctc	aaggaaagtaa	atctgtacag	gagaattggg	gagctgaaat	atttcccagt
82021	actgttttat	agtcttcttt	ttctttgttt	ttctttcttt	ttctcccaag	aaaactttta
82081	tggagtgtgt	tcaattatgt	atgcacgtat	gtgttcttcc	aaatggtgtg	tttctctgaa
82141	ccaacttcac	gtgtgggtag	cactgtgcta	gatactagct	gcataataaa	tgaattgtgaa
82201	agttgttgac	ataaggact	gatttttact	atctcattgac	cccttcgaac	atttatttgg
82261	ctgctatgat	tctatgcta	gaaagataca	tatatctaa	caaatacata	caatttttgg
82321	gagttctaga	tccttagaaa	gcctcccagg	ggtaacaaa	cagctagctg	ctgggpcgac
82381	tggctcatgc	ctataatgag	tactctctag	ctggggagat	cactggagcc	taggagtcca
82441	agaccagcct	agacaaact	gggagattcc	atttctacaa	taaaaaaatt	atctgggttg
82501	gggtatgctg	gcctgtagtg	ccagttactt	gggagctgta	ggctggagga	ccctctgaac
82561	ctggaaagct	gaggtctcag	tgaagcatga	ttatgtcact	gactccgaac	cagggaacaa
82621	gtgagactct	gtttcaaaaa	agaaaaccca	aaaaacaaa	aaacccaaca	gaactctgct
82681	ctaagaacct	agatttaaaa	aagcagcagt	gtaggctggg	caggtgtgct	catgcctctg
82741	atctcagcac	tttgggagcc	cgaggtgggc	agatcacttg	aggccaggag	cttaagacca
82801	gcctggcgaa	acactgtctc	tactaaaaaa	tacaaaaaatt	agtgtagtgt	ggtgatgcac
82861	gcgggtatct	ctagctactc	tggaggctga	ggcacaagaa	tcaattgaac	cggaaggttg
82921	gaggttgtag	tgaactgaga	tgtgccact	gcattccagc	actccaagct	tggtaatgag
82981	agactccctg	ttaaaaacaa	acaaaaacaa	acaaaaacaa	acaaaaccaa	gcctgtaatc
83041	ctgtgtgagt	aataaaaaag	gttttggctg	ggcgcgtggt	ctcatgcctg	taattccagc
83101	actttggggg	gctgaggcag	cagcagatca	ctgtgtcagg	agttttgagc	gcagcttgacc
83161	aacatggtga	aacctcgctc	ctactaaaaa	tacaaaaaatt	agcttgatgt	gggtgtgagc
83221	gcctgtgaag	ccagctgctc	aggaggctga	ggcaggagaa	ttgcttgaat	ctggggagcg
83281	gaggtctgac	tgaagcaaga	ttgtgccact	gcactccagc	ctgggtgaca	gagcagactc
83341	ccatctcaaa	aaaaaaaacaa	aaaaaaaacaa	aaaaaagctt	tgccaatgac	aatcatcagt
83401	tgtgtcagag	tgtctagaag	aaagtccacg	agtttcggcg	ggcatggttg	cgctgcctcc
83461	taattcccgac	actttggggg	gccgagatgg	gtggtacaca	aggctaggag	atcgagacca
83521	tctctgtcaa	cacggtgaaa	ccccgtctct	actaaaaata	caaaaaaaaa	ttagctYggcg
83581	ggggtggtg	gtgcctgtag	tcccagctac	tcaggagcct	gaggcaggag	aatgcgctga
83641	acccgggagg	tggagcttgc	aStgagccga	gatcgcgccca	ctcacgctcca	gcctggggtga
83701	cagagcaagaa	ctccgtctca	aaaagaaaaa	aaaaaagaaa	gtcacagagt	tgaccttaaa
83761	gctgtgggcaa	gatttggagt	gYgcagtaga	cagaaaggag	caggggtgag	gcacccaaag
83821	gagcaaggag	gcagacagca	tgaatagggt	gagggggcag	ctggaaaggag	gtcatgttca
83881	gtccaaagtct	cggttcggac	ctcaactgat	agattgacca	gggtttctcc	cagctgatac
83941	tgttgacatt	tggaaactgga	cagtctcttt	ttgtgaaggc	cagttctgtg	actgttaaga
84001	tgtccaacaa	cacacttgac	ctctgccacc	tagatgccag	tagcatctcc	caagtgtact
84061	caatcacaaa	tgtctccaga	cattgcccac	gatccccctg	gggacaaaga	ctccacctat
84121	ttagtatgac	gtgcacagca	tattttagct	tggcaatgtt	ttccaaagct	tagcatctca
84181	ggatacaact	tcacgaattt	tgccagatct	tgggtgcctcc	ttcttggcta	ttctcctgat
84241	ttatttattt	attttagggc	agaggctgag	gcaggagagat	cKcttgagcc	caggaggcag
84301	aggttgcagt	gagccgagat	caggccactg	cactttcagc	ctgggtgaca	catgagactc
84361	ccctctccaa	aaaaaaaacaa	aaaaaaaacaa	agtaaaagat	tacttctctc	cagatgttgg
84421	caaaaaaggt	ggagggccaa	gacRattggc	ctcttctctc	tacagagaggt	cccaacttat
84481	ccctgtgccca	tcatcatggt	agtagctctc	ggggcagatt	gcaagacccc	ctcttttcca
84541	ggagtgaatt	cccccttgcc	atgtgtcaaa	tccttctttg	cccttctttg	ctgatgagag
84601	gggtcatccc	acttcttgat	ctctcttctt	ggcaacagaa	gggtgatgtg	agatttaatt
84661	cttcaaaaaa	actaaaaaag	atgccaaagt	ccccagatgc	cgctccacgt	ggggccaagt
84721	ctctcttggg	tgcagagggt	cagtgacagc	aaagcgactt	ggatcgcttg	caggtgttcc
84781	ttctctctcta	attatagaaa	gtatttagct	attattccct	aatgtctcat	tgcagtttgg
84841	ttttagcatt	tttttttttt	tcctgacaaaa	ggctgtgttc	tgaatgagtc	tttttttttt
84901	tcgcgcttgg	ttcactgccca	taccagtgcc	tggctctctc	actggaacaa	tgcagagggt
84961	ggtaggttat	ttctggggaa	gccactttac	ctcatccaat	gtcttctcat	gcttccaggt
85021	tgctcaatag	gaaggggaat	aacatttact	gggcatctat	ttagtatgtg	ccaggcactg
85081	tgtcaagtgc	tcacgtaggt	tgtcttattc	aatccttgca	gcaaacctSa	gaggcagata

85141	taattattcc	Raatttatag	atgaagaatc	taagggtgaag	gtagctctcag	tcaagtattgtg
85201	gctgagtcag	aattccaaacc	cagggtctttg	tgtgtgaagt	ctttttttttga	ctttttttttga
85261	gcacatggtc	actctgtctca	ctcaggctgtg	agtgcmcgcg	tgcgaactctg	gctcactgga
85321	gcttccaaact	ccctgggtctca	aacaattctc	ctgctctcagc	ctcccaagta	gctgggacta
85381	cagggtgtgtg	ccacacacaca	tggtttaactc	tgttttatttt	tttaaatgaga	tggtctctca
85441	ccatgtgtgct	caggctgggtc	tgaactctct	gggtctaaagc	aattctctctg	ctggagtcag
85501	tctgggcaca	gtgggtgagc	cttgtaatcc	cttcccaaaa	tgtctgggatt	gtgagctact
85561	gtggccagac	tgaactccatg	tttctctctc	tcaaaaagttt	tctcagactt	ctgctgtata
85621	aaacatttgt	gttgggaact	cccgtctctc	tgcctcaaaa	gtgcttattt	ctaggtctcc
85681	tgttagtgtt	ggggttgtgt	ggggatgcct	atagagcgct	taaaaggagct	gtgttttcc
85741	tgttttgtgt	cttgacacat	gaaggaaattt	ccagtgtcat	ggttagcaagg	acgagggaga
85801	aagccatctc	gttctgggct	gtgcagtaaa	ggaatcctatg	ggaaacagat	ggcacattca
85861	ggggtttaca	tgcagaacag	aatattcaat	gaatggacta	cttagcaagc	ttatgggacg
85921	ttaaaaggac	caacaagcaa	tggggaaact	accgggtact	ggcagcagca	ggagccatcc
85981	accaccctca	ggtttaaagg	gacatgggtg	gagaatgggt	tctctgaagc	ccatttgacag
86041	ctgggggtat	gcgagaagca	tccagacagc	agctgtgaca	ggcagcagca	cgaccactct
86101	cccgaacccc	agcaggaagg	gcagcagcag	ctggtgtgtg	agttgtgtata	tatacacaga
86161	tccaagttct	tccctctctc	gccttcccc	ctgtctatgc	taccactggc	caaacccagc
86221	tagaagcagc	agagtagggg	agcctagaga	tgcagctctt	gcaggtctcag	ctctctgagc
86281	acaggcgagg	aaggaaaggtt	ggagagagca	ctctgggaaga	tacagagagt	tctcagcaga
86341	ggggctgact	cacctgtgcc	atccacagca	gtgtcttttt	ttttttNgaag	atagagactct
86401	tctctgttgc	ccaagtgtga	gtgcggtgat	gcaatctcag	ctcaactgaa	ttttgtctct
86461	ccgggttcaa	gcgatttctc	tccctcagcc	tccYgaatag	ctgggagctac	ataaagtcac
86521	caaccaacc	agctaattttt	tgtattttta	gtagagacgg	ggtttttgtta	tgttggtcag
86581	gctggtctca	aactcctggc	ctcaagtatg	ccaccctctc	tggcctccca	aagtgctaaag
86641	atttacaact	tgaagtcacca	cacctggccc	agtgctcaatt	tttttgtgtg	tgttgacata
86701	ctcttataga	agtgggacat	tgtttacagg	ctcaaaagat	catggaatgt	gagagtggga
86761	agtgccctta	gagctctacc	tgtaaaaattc	ctttacttta	taaaaggagat	ctctgagccc
86821	aggaagggga	aaggactggt	ccaaggtcag	tgtctaatct	gtggcctggy	caggacttga
86881	agccagcttg	cttaactctc	agtcctactc	gcttcttagt	tttgtagctgt	agtcacttga
86941	tggtttgttaa	taaaaggatta	ctactgagcc	acttgacctta	tttctcgttta	agacctcaga
87001	gtatagttta	ctttggcccc	agacatttgc	ctcaaatgat	tcttctctca	aatatactta
87061	tgtactgaggt	aggtctgccc	tttcagagtg	tgggctctgtg	ctcagggctc	atggggctga
87121	atccccagcc	tatctttaca	gaatggtctt	ggacaagtta	tttaaccctc	acccttattt
87181	tgtctcaactg	taaaatgagg	ataataatga	tgtctatctc	ataatctcatg	aggattatata
87241	aagaagatat	attttaaagca	cttagactag	catctgccaa	taatgagtcg	atagataaata
87301	acagtatatca	tgtggggtgt	gtgtgtatata	atatatatga	tgtaaacaaa	tatgtgtgtg
87361	tgtgtatgta	tatgcatgta	tgtgtatata	tatgcatgta	tgtgtgtgtg	tgtgtgtgtg
87421	tatatatata	tattctaccc	cccaaatag	tttagaattc	tgtttaaagaa	ttcacatggt
87481	tggctggggt	gtgggtggctc	atgtctgtaa	tccagcaact	tggggagact	aaggtgagca
87541	gatcatctga	ggctcaggagt	ttgagactag	cttgggccaa	atgggtaaat	ccctgttccc
87601	ctaaaagtga	aaaaattagc	cagggtgtgat	ggtagtgtgc	tgtagtccca	gctacttgca
87661	aggtctgagc	aggagaaagc	cttgaacctg	ggaagtggag	gttgacgtga	cccgaactgt
87721	cacactgagc	tccagcctgg	atgacaaagc	gcgactccat	ctcaaaaaaa	aaaaaaaaaa
87781	gaattctatgt	ggtttgctgg	gcaaggtggc	ttatgctctg	aaactccagca	cttagggagc
87841	ccaagcgagg	tggattgctt	gagctcgagg	gtttgagacc	cacctgggaa	acttgatgaa
87901	ataccactct	tatcaaaaat	acaaaaaaaa	aaaaattagc	tgggcatgtg	gggtgttgtt
87961	tgtgtctctc	gctacttggt	aggctgaggt	gggaggaatc	cttgagcccg	ggaggtggag
88021	gttgcagtga	accaagattg	tgccattgca	ctactccagc	ctggggcaga	gagaccacat
88081	ctaaaaaaaa	aaaaaaaaaa	aagaattcac	gtggctgtga	ttgagctcaag	gcattgttgt
88141	gtgtgctatc	agtcacagct	acttgggagg	ctgaggcgag	agaaacgctt	gaacctggga
88201	ggtggaggtt	gcagtgagcc	gacatcacac	actgcactcc	agcctggagt	acaaagcgag
88261	actccactca	atgctctgac	tttgggaggg	ttagtaggga	ggagtggctg	agcagcaggg
88321	tctcagatca	gctcaggcaa	catagggaga	cactgtcttt	gcataaaaaa	acaaaaaaat
88381	tctgggcatg	atgcaccacg	cctgtagctc	cagctactgt	ggagcgtgag	gtggggagat
88441	tacttgagcc	tgggaaggtca	aggctgtagt	gagtgcgcca	tgtatgccc	actgcattct
88501	agcctgggca	atacaacaag	accctgtctc	aaaaaaccaa	ccacccaacc	atacttttta
88561	cccttatgta	aatatatcag	tgttcttttt	aaacattttt	tctcttttaat	tttttttttt
88621	gtagagacaa	ggtctcaacct	atgttgccca	ggctggtctc	aaactcctaa	gctcactgaa
88681	accctctacc	ctgcctccc	aaagtctgg	gattacagggc	atgagcagca	gctcctggcc
88741	ttatcagctg	ttttccacat	ggaaattttc	ctttccctat	aaacagataag	aaactctgaag
88801	ggcttaggtg	atgtgccttt	cttgggactg	ctcagcagtc	agagcgagat	aaactcctgt
88861	tctcaggtg	ggcactgtgtg	ctcacactgt	taatccagcg	actttgagag	gctgagcgag
88921	cgggatcacc	tgaaggtcag	agtttgagac	cagctgggct	aaactgtgga	aaaccagctc

88981 cgactaaat acaaaaaat gccaggcatg gtggtgcatg cctgtaatcc cagttactcg  
89041 tggggctgag gcagggaat tacttgaact tggggagcgg tgggtgaactg gactgagctt  
89101 aaatcaaac tctccagccc tgggcaacag atgtgagact catctccaaa aaaaaaaag  
89161 aaatcaaac tctgactctt cacatcataa gctgctctgt ttggtctcatg gctgtaact  
89221 cagactcttg gggagctaaa gcggatggat catcttgaggc caggagtcttg agccagccct  
89281 ggcacaact gcgaaacccc gtctctacta aaatacaaaa aaatcaggcg gcggtgggtg  
89341 agctgtgtctg tgcctccagc tacttgagag gctgaggcag gagaatcgtc tgaacgggg  
89401 aggcagagt tgcagtgagc cggagtgggt cgggtgcaat cagactcgtg tgacagaaag  
89461 gactcgtctg tcaaaaaaaa caatacaaac acaacaacaa acaacaaggc aaagtgtgag  
89521 ctccacatta agcctgtacc ctgtgtgtaa catgggaaga caatgaagaag caatgactcg  
89581 agctttaattc atttatattc ggacataaag tccaccaaga taaaagaacat ctgtaccaaa  
89641 ctgccccttc cccacaaaat tcaagtattg ataagtattt ttaataaaat ttataaaat  
89701 aaacagttaga aggaggtaga gtgtgtctca gggagtgtg ggggtgagtg gggataggg  
89761 gtaattgggg atgacctctc tggagRgtaa cattagcaga gaacaaattg ccaggatgga  
89821 gcaagacatt gaagatatgg gggaaataga agtgaacagg ctgcccagat agcagccagc  
89881 ttggcgatt gagWacagaa aggaagctga tgcggttga gtggagtga caggggaaga  
89941 gaaaggtggg aaatgagctt ggaagagagg agtgagacca gcctctgagg agccaggaag  
90001 ctacagctga aggggtgggt tgcccctcca cactgtggg acactgtgga ccaggtgaac  
90061 gaggagcttg gaaaagaaaa agacacagag acaaaatata gagaagaaaa taaggggacc  
90121 cKggggacca ggtttcagca tatggaggat cccgccagcc tctgagttta cttagtattt  
90181 attgatctct cgtgggtgtt tctccgagag gggagtgtgt cagggtcaca agacataaga  
90241 gggagagagg tcagcagaca aacacgtgaa caaaggtctt tgcatacatg caaagttgat  
90301 ggtatgaagt ctgtgctttt agatatgcat acacataaac atctcaatgt ttacaaagc  
90361 agtatgtctg cccgcatgtc ccacctccag cctaaaggcg ctcttcccc atctcagtg  
90421 atggaacctt caatcggttt ttataccag agacttccat gccccaggga gggcaggaga  
90481 cagatgcctt ctctgtgtct caactgcaag aggcagctc tctcttata tctccctctc  
90541 tccagcagaga ccttttaccg gagtgtggct gggggagcgt cagggtcttt ccttcccagc  
90601 agggccatatt tcagactatc acatggggag aaaaacttga caatacctgt cttctcagg  
90661 cagagtgctc tgcggccttc cgcagttttt gtgtccctgt gtagcttga ttagggagtg  
90721 gtgatgact ctgaaggagca tgcgtccctc aagcatctgt ttaacaaagg aactcttga  
90781 cggcccttaa tccatttaac cctgagtttg acacagcaca tgtttcagag agccaggggt  
90841 taggggtgag tgcacagaat ctcaaggcag aagaattttt cttagtacct aacaaaaatg  
90901 agtctcccat gtctatttct ttctacacag acacagtaac aatctgatct ctcttgcttt  
90961 tcccacactc aggcaaaagag ttgtgatttt ctctcagtg tcatgagact cctctgaact  
91021 gcaacattca gcttccaaa cacttgacc accctcagtg atctccacaa gccctttgag  
91081 tgagtagaga ctgacaatag tatttaactc attttatcaa gacagaaact aaagctctgc  
91141 caaaattaga aaacctgctg tgagtcaacc agttcaaat agattctttg actcaaatc  
91201 ggggttttta acctcagctg tgcattgaaa tctcacagg agctcttcaa aatacctatg  
91261 acctcagctc attccagta gttgaagggg tgctagtggc agctatggcc cagggaagct  
91321 tttaagcttc tccaggtgat tttttttttt agactgggtt ttctctgctc acccagctg  
91381 gaatgcggtg gtgtgtcaat catgctcac tgaaccttg acctctgagg ctcaagtttg  
91441 ctcccagctg gaggctcctg agtagttggg attttaggca tgagcactgt tgcctggccc  
91501 ttattattat ttattttttt attttattt attttttgag atggagctct cctctgtcac  
91561 ccaggctgga gtcagtgac atgatctcg ctcactgcaa cctctgcctc cagctgtccc  
91621 gcgagtatcc tgccttagcc tctagagcag ctgggactac aggcgtgtga caccagcccc  
91681 agctaaagtgt tgtattttta gttagagcgg gttttacca tgttggcga gctgggtgtg  
91741 aattcctgac ctcaagtgt ctgcctgcc ttgtctccct aagtgtctgg attacaggtt  
91801 tgagccaact caccctggct aatgctccct ttttctgtt cctaaagcac ttgttcta  
91861 cctcaattct aacctcttc acattattg gaagtatca atctcttat cttaagctct  
91921 atacattgca aacttttagt ggtggggct gtgtattta actctatact acttaattg  
91981 ctgctgctga attatctttt tttttcttt ttagtgacc ggtttagctc tctctctgt  
92041 gccagctgg agtgagtggt caagacYcat caaagctcac tgcagcctg actactggg  
92101 ctcaagatcc tccctccctg gccctcagag tagctgggac tacaggatg cccacccatg  
92161 cccagctaat ttttaatttt tctgtaRtct catagggtct ggctatgttg ctccagctg  
92221 tctctaaact ccaagctcaa gcgattctcc tgccttgccc tcccaaaagt tgcgtgttac  
92281 agggataagc cactgtgctc agtcatgct gYctttcaat aaagacaagt ttgttcaatt  
92341 aatgtagttt aaattatca aagttttatt tacagttgt gatfatgga cttaataaac  
92401 ctaccctgga gatcataagg atgttttctt tctctttaa agttataatt ttgggaagct  
92461 atagtgaaga gaaaaaaaaa gttgtaatt taaaatgct taatttcccc tgtaagctct  
92521 taacctacct tgaatttagt ttgtcatagg ggaatcaaca ggttaactggt ttgtccagct  
92581 catgtactga atactctatc ttccctcact gaactataat tccagccata cctcatctta  
92641 tagatgcctg catgtttggg cctctgaggt gggttcaatt cccctctctg ctctccat  
92701 tccaatatca catggtctta attgtatag cttaataata catgttgaca tggaggggtg  
92761 cgtccttgaa acacattttt cttctcaag aatgttttgg ctatgcttg ctctaaagtg

92821 gtgKtttttt tttgtttttt gtattttttt ttttttttga gacaaaagtct cgcctgttca  
 92881 cccaggtctgg aattcagttg ttctgatcttg gctcactgca acattgttct cccgggttca  
 92941 aggtatctct tgccacagcc ttctgagtag ctgggattac aggtgtgca cccactgcct  
 93001 ggctaatctt ttgttttttag tagagatggg ttcttaacct atgttgcacg ctgtgtagta  
 93061 acctctgacc tcaactgact cactctgcctt agcctcccaa cgtctgtgga ttaaggcaca  
 93121 gagcctgacc gtctgcccttt gttttttcat acacactttt gatlagtttg ttacaggctca  
 93181 gaaaaataa tgttggaaa ttgattggaaa ctgcattgga ctgatagatt aatttcccca  
 93241 tgaataaaaa actcaagag atgtttatag aataaaca aaactgtgagc tgcgctggt  
 93301 ggctcacacc tgtaatccca gcaactttggg agggcaaggc cagagagatgg cttagagaaa  
 93361 aactctcgga ccagcctggg caatatagtg agaccacacc cccagctttt ttctgagaaa  
 93421 aattttaaaa aatttagcca ggcgtggcag tatgtgcttg tagtccacag tactccaag  
 93481 actgaggcag gagggtacct tgaggacag aggttgaggg tacagtgatg gagctctac  
 93541 actccagctt gggcaacaca gccagacctt gtctcaaaaa aaaaaagaga aatgtttttg  
 93601 atacatgtct tcaaatgtga tcttataagt caacatttga taatggagat cagtgaatga  
 93661 atgaataaag gccaaaagata actaagatta gctagcatat tatattttat tatttttta  
 93721 ttttgagcag agYcttgcct tgttgccca gctggagtgcc agtggcccaa tctcagctca  
 93781 ctgcaacctc cactctctgg gtccaagcaa ttctcccaac tcagctctccc gagtagttg  
 93841 gataacaggt gccgccact atgctgggct aattttttct tctatttttg tctcagctca  
 93901 agacagagct tcaacttgtt gcttagggctg gagtgagatg cgtgtatctg agctcactgc  
 93961 aactctcgcc tccggggttc aagcgattct cctgcctcag cttcctgagt agctgaattt  
 94021 ccagagctgc accaccacac caggctlaatt ttgtattttt tagcagagac agattttcact  
 94081 attgttgcga gctgggtctc aaattcccg gctcaagtgg tccgcccaac ttggcctccc  
 94141 aaggtgctgg gattacagga atgagccacc gcaaccggct taattttttt atttttagta  
 94201 gaaatgggag ttctactatgt cggcctggctt ggtctccaac ccttagctct cgtgtatctg  
 94261 cccgctcag cctccaaaag tgctgggatt acaggtgtga gctactgctg cctggcagtc  
 94321 cagcatattt attattattt ttgagacgga gtcttgcctg ttttctgctg cctcagctc  
 94381 aatggcgtag cctcggtcca ccgcaacttc tgctcccggt gttcttagtg cctcagctc  
 94441 ctgagtactg ggtgattacag gcatgcacca ctaggcctgg ctattttttg ctattattag  
 94501 tagggagcgg gtttctccat ttgtgtcagg ttggtctcaa acccccgacc tcaagtgatc  
 94561 tgcctacgtt ggctgctgg gattacaggg ctgagccacc gctggccgac cactaccaa  
 94621 ttgtcagtaa attctctctt cgctactagg ttccaatat ttattttaaa tctatggccc  
 94681 ggrtgagtggt cctcaggttg taactccagg actttgggaa gtaaggcgac ttggtattgc  
 94741 tgagtccagg agttcaagac cagcctgggg aaaaatggca aaccccgcca cctcctaaaa  
 94801 tacaataaag tgcaggggcg ttgtagcaca ctctgttaat cccagctact caggaggctg  
 94861 aggcaggagga atcacttgaa cccgggaggg ggaggttgca gtgaggcgag atgggccact  
 94921 gcactccagc ttgggcgaca gagtgaact ctgtctcaaa atgaatgaat gaatgaatga  
 94981 atgaatgaat aaaaactacg aggtcggggc cagtggctca cctcttgtaat cccagcaact  
 95041 tggcggggcg gatcatttga ggtcaggagt ttgagaccgg cttggccaac atggtgaac  
 95101 cccactctcta aaaaaaatc aaaaatttag ttactgtctt gggcgacaga gcgagactca  
 95161 gtctcaaaaa aaaaaatct acaaatccac ttactcttct attaaacag aacctgggac  
 95221 tagtttagaa tgtcttgaag aaaaatttct agtaatatc caagtccgct gccctggcga  
 95281 agcaacaagg ccactgaagc gagtgaagcc gagaccctgt tgctccaca cctgtgcaac  
 95341 cgggatcgcg cgtgggggaa ccaggaaact caagtcggcg caggcccgcg ccgctctgga  
 95401 tctactacgc cagtcagggt ggcgtgcccc gcaaggttcc cagctaatct cgcctaccg  
 95461 ggttgccggc ggaagccggg cgcgcggct ctgcttccct cggggtgagt aggggagct  
 95521 cgcggcgagc cggcgagag ctcggggcga ccagcctggc agggccgac cgcggcgga  
 95581 cgcggagga accgcatgg acacgcggg agtcggcctc tgcctcgggt ttgcagcccg  
 95641 cctggcgccc gggagggaca gcccatgac cttgtcaact gggcccgagc gggcccgag  
 95701 tggctcgggt gactcgact gccccggccc cctgcagtga cttctccagt ttggcacacc  
 95761 agtgggcttt catgtgacct gtccacttca gttggcagat gttggcagga cccgagtggg  
 95821 cagagcttga ggaaggccgg aggtggctca cccagacgac ttaccacctt ggtggccctg  
 95881 tgaatttcga gccgaagaca gctggagctg gagcgtggtc tggcgagag tggggcgga  
 95941 tgcccatctc ctggccctcc acccggtctt gcgaaagtgc cccagacgca tgcttatctc  
 96001 cagcatcttc cagggccctt ggcctccaga tcaacagta tcatatgtgt cttaccttca  
 96061 agtgtgtgtg gagagcaacc tttagataag acactagaa cagccctctc ctaggtttaa  
 96121 tgggcatctt ctttctacta acaagaaggt cattcaaca gattttat tg ttaccttca  
 96181 tgaactctgg cacttagtag gtgctcaaga aatatlttga gaggttgaat taacttctc  
 96241 ttactctggt cttcatatgt ttattttaat aattttcttg gggctggaat tgggtcgta  
 96301 tgcctgtaat cccagcactt tggagggcca agatggggag atgtctgtag ggcagagatt  
 96361 caacaccagc ctgggcacaa tagtgagact cctctgtga ttaaaagaaa caagatgaaa  
 96421 taaaataaaa aatttttcta ctacagtggt tctcaaaact tagcatgca cagagtcaac  
 96481 ttgagagctc gtaacaacac acagactgg acccactccc caaagtctct gatcttctga  
 96541 gtctggtagt agSaccaga attggcgttt ctaactatga ccaagtgat caggttgcgt  
 96601 tgggtctggg gaccatacgt tgagtaccac tgctttatag gaaaaaaaat cctgtgaact

```

96661   taaagcacta   tctcttcaca   cctgggtacg   ggtctttctg   tctttatttg   ccagttttac
96721   ttttttggtg   attgtaaata   tgtgcataatt   actttgtaca   ctgttaactc   cctccatatt
96781   aatttatata   cagattttta   tgtagtctat   acttgctatt   tttcctgggt   acatcatatt
96841   ctaccttatt   gggctattct   ttgcttggtc   atttctcaac   tggcctcagg   gatgtcgat
96901   ttgatctgta   ctagaagagt   gtgcatactc   ttactttctg   Miatgcagtc   attttcttgt
96961   cagacaaaag   acctgtaaga   atattagttg   gggccgggtg   cggtaggtca   cacctgtaat
97021   cccagcacit   tgggaggctg   aggtgggcga   atcacttgag   gtccaggcgtt   caatgccagc
97081   ctggccaaca   tggtgaaaac   ccactctctac   caaaaataca   aaaaatttac   ctgggcgtag
97141   tgggtgtgtc

```

[0256] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the aspects which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0257] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.



What is claimed is:

1. A method for identifying a subject at risk of osteoarthritis, which comprises detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variations are detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c);

whereby the presence of the polymorphic variation is indicative of the subject being at risk of osteoarthritis.

2. The method of claim 1, which further comprises obtaining the nucleic acid sample from the subject.

3. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 31118000 to 31129000 in human genomic DNA.

4. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 selected from the group consisting of 247, 1535, 2386, 6440, 9133, 9143, 9471, 13150, 13717, 14466, 15769, 16870, 18545, 18749, 19123, 20736, 21038, 21046, 21050, 21056, 21706, 23170, 25028, 27871, 28070, 31717, 32019, 32318, 33080, 33101, 34236, 34285, 34818, 35168, 37981, 38113, 38117, 38481, 38615, 38944, 39288, 41385, 42136, 42185, 42353, 42434, 44580, 44675, 45739, 46439, 47457, 47735, 50319, 50708, 51185, 53002, 53064, 53637, 55274, 55825, 55986, 56684, 57653, 57659, 57692, 57775, 61313, 61431, 61699, 62906, 63619, 64664, 68452, 69665, 69681, 70091, 74637, 74760, 76523, 78559, 79549, 79882, 81339, 81681, 81696, 83517, 85431, 86332, 87358, 87725, 89052, 90020, 90231, 90284, 90447, 90601, 90724, 92559, 95176, 95195 and 96822.

5. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 selected from the group consisting of 13150, 21046, 23170, 25028, 44580, 62906, 64664 and 83517.

6. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 36914000 to 36931000 in human genomic DNA.

7. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 2 selected from the group consisting of 211, 7217, 7895, 13308, 14279, 17026, 18271, 20417, 21843, 22069, 22145, 22519, 22539, 23236, 23256, 23402, 23499, 23620, 23871, 24136, 25427, 25866, 26541, 26576, 26689, 26720, 27113, 27164, 27186, 28341, 29160, 29844, 30665, 30830, 31061, 31523, 32326, 32346, 32358, 34909, 34975, 35066, 35096, 35375, 36304, 36712, 36770, 37342, 37412, 37884, 38077, 38300, 38301, 41189, 44408, 44493, 44571, 44670, 45219, 45258, 47261, 48473, 48771, 55292, 56479, 56747, 60620, 60688, 61058, 61129, 61577, 61961, 63351, 63926, 65798, 66043, 66044, 66246, 66318, 66547, 71238, 71283, 71492, 72274, 73762, 74209, 75284, 77347, 77589, 78096, 78606, 78862, 79135, 79146, 79456, 79609, 80086, 80119, 80766, 81110, 81269, 81668, 82433, 82559, 83298, 83821, 84121, 84147, 84543, 84554, 84691, 84727, 85678, 86699, 86700, 86792, 86832, 87045, 87140, 87365, 88342, 88498, 88589, 95502, 96968, 97448, 97568 and 98724.

8. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 2 selected from the group consisting of 23236, 32358, 47261, 48771, 55292, 60688, 72274, 74209, 77589, 79135, 79456, 79609, 80119, 80766, 81110, 82433, 84121, 84147, 85678, 86699, 86832, 87140 and 88589.

9. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 76196500 to 76221500 in human genomic DNA.

10. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 3 selected from the group consisting of 218, 1440, 1442, 2611, 4317, 4724, 4788, 5202, 5780, 5974, 6644, 7430, 7938, 8095, 8183, 8312, 8352, 9348, 9378, 9617, 9727, 9834, 9899, 10211, 10377, 10695, 10729, 10730, 11433, 11951, 12697, 12982, 14419, 14501, 14983, 15280, 15475, 15888, 15976, 16307, 16442, 17255, 18948, 19435, 19753, 20021, 20022, 20503, 20590, 21804, 21919, 21990, 22412, 22536, 23432, 23468, 23772, 24325, 24773, 26274, 27440, 28561, 30071, 31764, 33008, 35310, 35460, 37112, 37285, 37747, 38057, 38859, 38860, 39525, 40216, 40281, 41453, 42091, 42513, 42935, 42985, 43003, 43281, 43716, 43866, 44234, 44596, 44871, 45005, 45282, 47178, 47816, 47887, 48134, 48135, 48276, 48400, 48798, 48803, 49146, 49969, 51059, 51064, 53285, 54560, 54748, 54785, 55102, 55644, 55705, 55841, 56623, 56825, 56827, 56892, 59150, 59958, 60231, 60524, 61871, 62226, 63230, 63468, 63787, 65732, 65989, 68832, 69904, 70365, 70886, 73088, 73103, 75934, 75966, 76273, 77943, 78466, 78861, 78872, 79836, 80908, 81509, 83576, 83662, 83782, 84282, 84444, 85129,

85151, 85296, 85809, 86387, 86494, 89786, 89894, 90122, 92067, 92187, 92312, 92824, 93733, 96553 and 96941.

11. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 3 selected from the group consisting of 4788, 8312, 9378, 9727, 9899, 10211, 27440, 40216, 40281, 42091, 43866, 48803, 51059, 55644, 56623, 73103, 78872, 79836, 85129, 92824 and 96941.

12. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in Table A.

13. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in linkage disequilibrium with one or more positions in claim 4, 7, 10 or 12.

14. The method of claim 1, wherein detecting the presence or absence of the one or more polymorphic variations comprises:

hybridizing an oligonucleotide to the nucleic acid sample, wherein the oligonucleotide is complementary to a nucleotide sequence in the nucleic acid and hybridizes to a region adjacent to the polymorphic variation;

extending the oligonucleotide in the presence of one or more nucleotides, yielding extension products; and

detecting the presence or absence of a polymorphic variation in the extension products.

15. The method of claim 1, wherein the subject is a human.

16. The method of claim 15, wherein the subject is a human female.

17. The method of claim 15, wherein the subject is a human male.

18. A method for identifying a polymorphic variation associated with osteoarthritis proximal to an incident polymorphic variation associated with osteoarthritis, which comprises:

identifying a polymorphic variation proximal to the incident polymorphic variation associated with osteoarthritis, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation;

determining the presence or absence of an association of the proximal polymorphic variant with osteoarthritis.

19. The method of claim 18, wherein the incident polymorphic variation is at one or more positions in claim 4, 7, 10 or 12.

20. The method of claim 18, wherein the proximal polymorphic variation is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the incident polymorphic variation.

21. The method of claim 18, which further comprises determining whether the proximal polymorphic variation is in linkage disequilibrium with the incident polymorphic variation.

22. The method of claim 18, which further comprises identifying a second polymorphic variation proximal to the identified proximal polymorphic variation associated with osteoarthritis and determining if the second proximal polymorphic variation is associated with osteoarthritis.

23. The method of claim 22, wherein the second proximal polymorphic variant is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the proximal polymorphic variation associated with osteoarthritis.

24. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

wherein the nucleotide sequence comprises a polymorphic variation associated with osteoarthritis selected from the group consisting of in SEQ ID NO: 1 a guanine at position 13150, a thymine at position 21046, an adenine at position 23170, an adenine at position 25028, a guanine at position 44580, a guanine at position 62906, a cytosine at position 64664 and a cytosine at position 83517; in SEQ ID NO: 2 an adenine at position 23236, a cytosine at position 32358, a guanine at position 47261, a guanine at position 48771, a cytosine at position 55292, an adenine at position 60688, a guanine at position 72274, a guanine at position 74209, a cytosine at position 77589, an adenine at position 79135, a thymine at position 79456, an adenine at position 79609, an adenine at position 80119, a cytosine at position 80766, an adenine at position 81110, a cytosine at position 82433, a cytosine at position 84121, a thymine at position 84147, a cytosine at position 85678, a thymine at position 86699, an adenine at position 86832, a guanine at position 87140 and an adenine at position 88589; in SEQ ID NO: 3 a guanine at position 4788, a thymine at position 8312, a deletion at position 9378, a cytosine at position 9727, a guanine at position 9899, a cytosine at position 10211, a guanine at position 27440, a guanine at position 40216, a cytosine at position 40281, an adenine at position 42091, a guanine at position 43866, an adenine at position 48803, an adenine at position 51059, an adenine at position 55644, a cytosine at position 56623, a cytosine at position 73103, an adenine at position 78872, a guanine at position 79836, a cytosine at position 85129, a guanine at position 92824 and an adenine at position 96941; and an allele associated with osteoporosis in Table A for positions rs552, rs12904, rs2282146, rs734784, rs1042164, rs749670, rs955592, rs1143016, rs755248, rs1055055, rs835409, rs927663, rs8162, rs831038, rs33079, rs1710880, rs4472029, rs799570, rs1282730, rs1518875, rs1568694, rs905042, rs1957723, rs794018, rs707723, rs893861, rs1914903, rs2062232, rs26609, rs1370987, rs1012414, rs435903, rs1248, rs703508, rs226465, rs241448, rs763155, rs1040461, rs462832, rs804194, rs1022646, rs1569112, rs805623, rs1019850, rs1599931, AA, rs279941, rs1062230, rs1859911, rs1477261, rs1191119, rs657780, rs1393890, rs1478714, rs868213, rs690115, rs1465501, rs899173, rs10477, rs926393, rs465271, rs13847 and rs738658.

25. An oligonucleotide comprising a nucleotide sequence complementary to a portion of the nucleotide sequence of (a), (b), (c), or (d) in claim 24, wherein the 3' end of the oligonucleotide is adjacent to a polymorphic variation associated with osteoarthritis.

26. A microarray comprising an isolated nucleic acid of claim 24 linked to a solid support.

27. An isolated polypeptide encoded by the isolated nucleic acid sequence of claim 24.

28. A method for identifying a candidate therapeutic for treating osteoarthritis, which comprises:  
(a) introducing a test molecule to a system which comprises a nucleic acid comprising a nucleotide sequence selected from the group consisting of:

- (i) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (ii) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (iv) a fragment of a nucleotide sequence of (a), (b), or (c); or

introducing a test molecule to a system which comprises a protein encoded by a nucleotide sequence of (i), (ii), (iii), or (iv); and

(b) determining the presence or absence of an interaction between the test molecule and the nucleic acid or protein,

whereby the presence of an interaction between the test molecule and the nucleic acid or protein identifies the test molecule as a candidate therapeutic for treating osteoarthritis.

29. The method of claim 28, wherein the system is an animal.

30. The method of claim 28, wherein the system is a cell.

31. The method of claim 28, wherein the nucleotide sequence comprises one or more polymorphic variations associated with osteoarthritis.

32. The method of claim 31, wherein the one or more polymorphic variations associated with osteoarthritis are at one or more positions in claim 4, 7, 10 or 12.

33. A method for treating osteoarthritis in a subject, which comprises contacting one or more cells of a subject in need thereof with a nucleic acid, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c); and  
(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);  
whereby contacting the one or more cells of the subject with the nucleic acid treats the  
osteoarthritis in the subject.

34. The method of claim 33, wherein the nucleic acid is RNA or PNA.

35. The method of claim 34, wherein the nucleic acid is duplex RNA.

36. A method for treating osteoarthritis in a subject, which comprises contacting one or more  
cells of a subject in need thereof with a protein, wherein the protein is encoded by a nucleotide sequence  
which comprises a polynucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;  
(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in  
SEQ ID NO: 1-3 or referenced in Table A;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the  
amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c);  
whereby contacting the one or more cells of the subject with the protein treats the osteoarthritis  
in the subject.

37. A method for treating osteoarthritis in a subject, which comprises:  
detecting the presence or absence of one or more polymorphic variations associated with  
osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variation are  
detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;  
(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in  
SEQ ID NO: 1-3 or referenced in Table A;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the  
amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic  
variation; and

administering an osteoarthritis treatment to a subject in need thereof based upon the presence or  
absence of the one or more polymorphic variations in the nucleic acid sample.

38. The method of claim 37, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 12.

39. The method of claim 37, wherein the treatment is selected from the group consisting of administering a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondroitin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

40. A method for detecting or preventing osteoarthritis in a subject, which comprises:  
detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
  - (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
  - (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
  - (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and
- administering an osteoarthritis prevention or detection procedure to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

41. The method of claim 40, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 12.

42. The method of claim 40, wherein the osteoarthritis prevention is selected from the group consisting of administering a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondroitin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

43. A method of targeting information for preventing or treating osteoarthritis to a subject in need thereof, which comprises:



detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

directing information for preventing or treating osteoarthritis to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

44. The method of claim 43, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 12.

45. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and an antibody that specifically binds to a protein, polypeptide or peptide encoded by a nucleotide sequence identical to or 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A.

46. The composition of claim 45, wherein the antibody specifically binds to an epitope comprising an amino acid encoded by rs734784, rs1042164, rs749670, rs955592 and rs1040461.

47. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and a RNA, DNA, PNA or ribozyme molecule comprising a nucleotide sequence identical to or 90% or more identical to a portion of a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A.

48. The composition of claim 47, wherein the RNA molecule is a short inhibitory RNA molecule.

Abstract of the Disclosure

Provided herein are methods for identifying a risk of osteoarthritis in a subject, reagents and kits for carrying out the methods, methods for identifying candidate therapeutics for treating osteoarthritis, and therapeutic and preventative methods applicable to osteoarthritis. These embodiments are based upon an analysis of polymorphic variations in nucleotide sequences within the human genome.

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

## **IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

FIGURE 1A

KIAA0296 – DISCOVERY P-VALUES (female only)

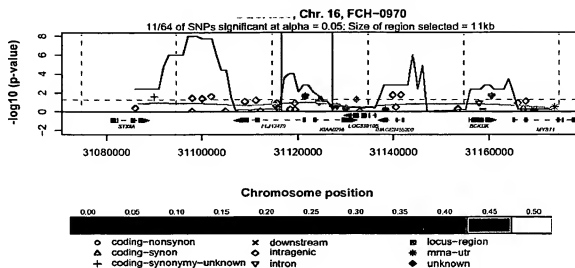


FIGURE 1B

CHROM 4 – DISCOVERY P-VALUES (female only)

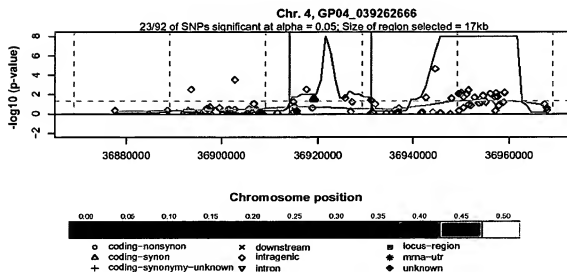
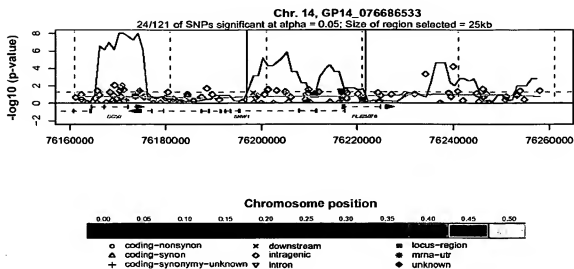


FIGURE 1C

SNW1 – DISCOVERY P-VALUES (female only)



## **Application Data Sheet**

### **Application Information**

Application Type::	Provisional
Subject Matter::	Utility
Suggested Group Art Unit::	Not Yet Assigned
CD-ROM or CD-R?::	None
Sequence submission?::	None
Computer Readable Form (CRF)?::	No
Title::	METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF
Attorney Docket Number::	524593008600
Request for Early Publication?::	No
Request for Non-Publication?::	No
Total Drawing Sheets::	2
Small Entity?::	Yes
Petition included?::	No
Secrecy Order in Parent Appl.?::	No

### **Applicant Information**

Applicant Authority Type::	Inventor
Primary Citizenship Country::	US
Status::	Full Capacity
Given Name::	Steven
Family Name::	MAH
City of Residence::	San Diego
State or Province of Residence::	CA
Country of Residence::	US
Street of mailing address::	12820 Via Nieve #74
City of mailing address::	San Diego
State or Province of mailing address::	CA

Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor  
Primary Citizenship Country:: Germany  
Status:: Full Capacity  
Given Name:: Andreas  
Family Name:: BRAUN  
City of Residence:: San Diego  
State or Province of Residence:: CA  
Country of Residence:: US  
Street of mailing address:: 3935 Lago Di Grata Circle  
City of mailing address:: San Diego  
State or Province of mailing address:: CA  
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor  
Primary Citizenship Country:: Germany  
Status:: Full Capacity  
Given Name:: Stefan  
Middle Name:: M.  
Family Name:: KAMMERER  
City of Residence:: San Diego  
State or Province of Residence:: CA  
Country of Residence:: US  
Street of mailing address:: 3825 Elijah Court, Unit 334  
City of mailing address:: San Diego  
State or Province of mailing address:: CA  
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor  
Primary Citizenship Country:: US  
Status:: Full Capacity

Given Name:: Matthew  
Middle Name:: Roberts  
Family Name:: NELSON  
City of Residence:: San Marcos  
State or Province of Residence:: CA  
Country of Residence:: US  
Street of mailing address:: 1250 Calle Prospero  
City of mailing address:: San Marcos  
State or Province of mailing address:: CA  
Postal or Zip Code of mailing address:: 92069

Applicant Authority Type:: Inventor  
Primary Citizenship Country:: Sweden  
Status:: Full Capacity  
Given Name:: Rikard  
Middle Name:: Henry  
Family Name:: RENELAND  
City of Residence:: San Diego  
State or Province of Residence:: CA  
Country of Residence:: US  
Street of mailing address:: 7555 Charmant Drive, #1114  
City of mailing address:: San Diego  
State or Province of mailing address:: CA  
Postal or Zip Code of mailing address:: 92122

Applicant Authority Type:: Inventor  
Primary Citizenship Country:: United Kingdom  
Status:: Full Capacity  
Given Name:: Maria  
Middle Name:: L.  
Family Name:: LANGDOWN  
City of Residence:: San Diego



State or Province of Residence:: CA  
Country of Residence:: US  
Street of mailing address:: 3701 Yosemite Street  
City of mailing address:: San Diego  
State or Province of mailing address:: CA  
Postal or Zip Code of mailing address:: 92109

**Correspondence Information**

Correspondence Customer Number:: 25225

**Representative Information**

Representative Customer Number:: 25225